M.M. POSTNIKOV

LEÇONS DE GÉOMÉTRIE

3e semestre

VARIÉTÉS DIFFÉRENTIABLES

M. POSTNIKOV

LEÇONS DE GÉOMÉTRIE

VARIÉTÉS DIFFÉRENTIABLES

ÉDITIONS MIR MOSCOU

TABLE DES MATIÈRES

Préface 9

Leçon première 14

Lignes simples du plan. — Définition d'une ligne par une équation. — Théorème de Whitney. — Courbes de Jordan. — Courbes régulières et différentiables. — Courbes non paramétrées. — Paramètre naturel.

Leçon 2 30

Courbes dans le plan. — Formules de Frenet pour une courbe gauche. — Projections d'une courbe sur les plans de coordonnées du repère mobile. — Formules de Frenet pour une courbe dans un espace à n dimensions. — Existence et unicité d'une courbe de courbures données.

Leçon 3 42

Surfaces élémentaires et leurs paramétrisations. — Exemples de surfaces. — Plan tangent et sous-espace tangent. — Applications différentiables de surfaces et leurs différentielles. — Difféomorphismes de surfaces. — Première forme quadratique d'une surface. — Isométries. — Premier paramètre différentiel de Beltrami. — Exemples de calcul des premières formes quadratiques. — Surfaces développables.

Leçon 4 67

Vecteur normal. — La surface, graphique d'une fonction. — Sections normales. — Deuxième forme quadratique d'une surface. — Indicatrice de Dupin. — Courbures principales totale et moyenne. — Deuxième forme quadratique d'un graphique. — Surfaces réglées de courbure nulle. — Surfaces de révolution.

Leçon 5 84

Formules de dérivation de Weingarten. — Coefficients de connexité. — Théorème de Gauss. — Formule explicite de la courbure totale. — Conditions nécessaires et suffisantes d'isométrie. — Surfaces de courbure constante.

Leçon 6 93

Remarques préliminaires. — Sous-ensembles ouverts de l'espace $\mathbb{R}^n$ et leurs difféomorphismes. — Cartes et atlas. — Atlas maximaux. — Variétés différentiables. — Exemples de variétés différentiables.

Leçon 7 107

Topologie d'une variété différentiable. — Sous-variétés ouvertes. — Points intérieurs et voisinages. — Homéomorphismes. — Premier axiome de dénombrabilité et euclidicité locale. — Deuxième axiome de dénombrabilité. — Variétés non séparées. — Structures différentiables sur un espace topologique. — Variétés topologiques. — Variétés 0-dimensionnelles. — Catégorie TOP. — Catégorie DIFF. — Transport de différentiabilité.

Leçon 8 120

Invariance topologique de la dimension des variétés. — Dimension par recouvrements. — Espaces compacts. — Lemme de Lebesgue. — Majoration de la dimension des sous-ensembles compacts de $\mathbf{R}^n$. — Monotonie de la dimension. — Ensembles fermés. — Monotonie de la dimension par ensembles fermés. — Produit direct d'espaces topologiques. — Compacité du produit direct d'espaces compacts.

Leçon 9 132

Théorème du tambour. — Théorème de Brouwer du point fixe. — Théorème des cloisons dans le cube. — Espaces normaux et complètement normaux. — Prolongement des cloisons. — Théorème de Lebesgue de recouvrements du cube. — Minoration de la dimension du cube.

Leçon 10 144

Ordinaux. — Topologie de l'ordre dans les ensembles d'ordinaux. — Espaces de dimension 0. — Exemple de Tykhonov. — Produit tykhonovien d'espaces topologiques. — Filtres. — Ensembles centrés d'ensembles. — Ultrafiltres. — Critère de compacité. — Théorème de Tykhonov.

Leçon 11 158

Structure différentiable sur un espace affine. — Variété de matrices de rang donné. — Variétés de Stiefel. — Séries de matrices. — Exponentielle d'une matrice. — Logarithme d'une matrice. — Matrices orthogonales et J-orthogonales. — Groupes de Lie de matrices. — Groupes de matrices J-orthogonales. — Matrices unitaires et J-unitaires. — Groupes de Lie de matrices complexes. — Variétés $\mathbf{C}$-analytiques. — Espaces connexes par arcs. — Espaces connexes. — Coïncidence de la connexité et de la connexité par arcs pour les variétés. — Chemins différentiables et différentiables par morceaux. — Variétés connexes ne satisfaisant pas au deuxième axiome de dénombrabilité.

Leçon 12 179

Vecteurs tangents à une variété différentiable. — Dérivées de fonctions holomorphes. — Vecteurs tangents à des variétés $\mathbf{C}$-analytiques. — Différentielle d'une application différentiable. — Règle en chaîne. — Gradient d'une fonction différentiable. — Théorème des applications étales. — Théorème de changement des coordonnées locales. — Applications localement planes.

Leçon 13 196

Démonstration du théorème des applications localement planes. — Immersion et submersion. — Sous-variétés d'une variété différentiable. — Sous-espace tangent à une sous-variété. — Définition lo-

cale d'une sous-variété. — Unicité de la structure d'une sous-variété. — Cas des sous-variétés plongées. — Théorème de l'image réciproque d'une valeur régulière. —Solutions de systèmes d'équations. — Le groupe SL (n) traité comme une sous-variété.

Leçon 14 209

Théorème de plongement. — Retour sur les ensembles compacts. — Fonctions d'Urysohn. — Démonstration du théorème de plongement. — Variétés satisfaisant au deuxième axiome de dénombrabilité. — Ensembles clairsemés et ensembles maigres. — Ensembles négligeables.

Leçon 15 221

Théorème de Sard. — Partie analytique de la démonstration du théorème de Sard. — Produit direct de variétés. — Variété de vecteurs tangents. — Démonstration du théorème de plongement de Whitney.

Leçon 16 233

Tenseurs. — Champs de tenseurs. — Champs de vecteurs et dérivation. — Algèbre de Lie de champs de vecteurs.

Leçon 17 247

Courbes intégrales de champs de vecteurs. — Champs de vecteurs et flots. — Transfert de champs de tenseurs par des difféomorphismes. — Dérivée de Lie d'un champ de tenseurs.

Leçon 18 258

Formes différentielles linéaires. — Formes différentielles de degré quelconque. — Formes différentielles en tant que fonctionnelles de champs de vecteurs. — Produit intérieur d'un champ de vecteurs par une forme différentielle. — Transfert d'une forme différentielle par une application différentiable.

Leçon 19 270

Différentielle extérieure d'une forme différentielle. — Dérivée de Lie d'une forme différentielle.

Leçon 20 280

Complexe de de Rham et groupes de cohomologie d'une variété différentiable. — Groupe $H^0\mathcal{X}$. — Lemme de Poincaré. — Groupe $H^1\mathbb{S}^2$. — Groupe $H^1\mathbb{S}^1$. — Calcul du groupe $H^1\mathbb{S}^1$ à l'aide d'intégrales. — Groupe $H^2\mathbb{S}^2$. — Groupes $H^1\mathbb{S}^n$ pour $n \geqslant 2$. — Groupes $H^m\mathbb{S}^n$, $m < n$. — Groupes $H^n\mathbb{S}^n$.

Leçon 21 303

Schémas simpliciaux et leurs réalisations géométriques. — Groupes de cohomologie de schémas simpliciaux. — Complexe double d'un recouvrement. — Groupes de cohomologie d'un complexe double. — Complexes doubles bordés. — Homomorphismes bords. — Complexes acycliques. — Acyclicité par rapport aux lignes pour $p = 0$.

Leçon 22 316

Acyclicité par rapport aux lignes d'un complexe double d'un recouvrement numérotable. — Acyclicité par rapport aux colonnes d'un complexe double d'un recouvrement de Leray. — Théorème de de

Rham-Leray. — Généralisation. — Groupes $E_2^{p,q}$. — Groupes $F^{p,q}$. — Groupe adjoint à un groupe gradué à filtration.

Leçon 23 329
Groupes $E_r^{p,q}$. — Suites spectrales. — Suite spectrale d'un complexe double. — Suite spectrale d'un recouvrement.

Leçon 24 342
Espaces topologiques dénombrables à l'infini et paracompacts. — Variétés paracompactes. — Intégrales dans $\mathbb{R}^n$. — Ensembles cubables et densités sur des variétés arbitraires. — Intégration des densités.

Leçon 25 355
Variétés orientables. — Intégration des formes. — Lemme de Poincaré pour formes à support borné. — Le groupe $H^n_{s.b}\mathcal{V}$. — Cas d'une variété orientable.

Leçon 26 360
Degré d'une application différentiable propre. — Nombre algébrique des images réciproques d'une valeur régulière. — Invariance du degré par une homotopie différentiable. — Démonstration du théorème du tambour. — Invariance du degré par une homotopie quelconque.

Leçon 27 378
Domaines à frontière régulière. — Théorème de Stokes. — Formules de Gauss-Ostrogradski, Green et Newton-Leibniz. — Variétés à bord. — Points intérieurs et points du bord. — ∂-sous-variétés plongées. — Théorème de Stokes pour variétés à bord et ∂-sous-variétés. — Théorème de Stokes pour intégrales de surface. — Théorème de Stokes pour sous-variétés singulières. — Intégrales curvilignes de seconde espèce.

Leçon 28 400
Opérateurs d'analyse vectorielle. — Corollaires de l'identité $d \circ d = 0$. — Corollaires de la formule de dérivation d'un produit. — Opérateurs de Laplace et de Beltrami. — Flux d'un champ de vecteurs. — Formule de Gauss-Ostrogradski pour la divergence et formules de Green. — La divergence en tant que densité de sources. — Formule de Stokes pour la circulation. — Formule de Gauss-Ostrogradski pour le rotationnel. — Formule généralisée de Gauss-Ostrogradski.

Leçon 29 410
Périodes des formes différentielles. — Simplexes, chaînes, cycles et frontières singuliers. — Théorème de Stokes pour intégrales le long de chaînes. — Groupes d'homologie singulière. — Théorème de de Rham. — Groupes de cohomologie d'un complexe de chaînes. — Groupes de cohomologie singulière.

Index 427

PRÉFACE

La géométrie a été et reste le parent pauvre des programmes de la faculté de mécanique et de mathématiques (« Mekhmat ») de l'Université de Moscou. Dans aucun programme des cinquante dernières années on ne trouvera la moindre trace ni des bases de la géométrie, ni des courbes algébriques, ni des groupes de transformations, ni même enfin de la géométrie projective (exception faite de quelques chapitres qui sont inclus dans le cours de géométrie analytique du premier semestre et qui ne sont lus que par un heureux concours de circonstances et parfois même escamotés dans l'indifférence la plus absolue). On peut très bien être diplômé de « Mekhmat » sans avoir la moindre notion de la géométrie de Lobatchevski, sans connaître les idées de Caley-Klein sur les fondements de la géométrie, les propriétés des courbes algébriques et des groupes de Lie!

L'implantation des méthodes géométriques vers les années 70 qui a eu pour effet de surcharger le cours d'analyse mathématique par des éléments de géométrie a inspiré un nouveau cours intitulé « Variétés différentiables et géométrie différentielle » qui était dispensé une fois par semaine et dont l'objectif était de décharger les professeurs d'analyse de l'exposition de cette discipline « étrangère ». Mais la conception de ce cours ne fut pas très heureuse et de plus la jonction avec les cours parallèles d'analyse et de théorie des équations différentielles fut mal assurée. Le résultat fut piteux et frisa l'anecdote: l'intégration des formes différentielles sur des variétés et la formule de Stokes furent exposées simultanément et dans le cours d'analyse et dans celui de géométrie!

L'enseignement du cours de géométrie au troisième semestre est entré en contradiction avec le rôle généralisateur et unificateur des notions géométriques en mathématiques modernes, rôle qui ne peut être mis en évidence qu'après l'assimilation des principales disciplines de l'analyse.

Tout ceci a conduit à reléguer le cours de géométrie en troisième année (cinquième et sixième semestres). Mais on s'aperçut immédiatement que cette solution avait aussi ses défauts.

La clef de voûte de tout cours de géométrie est la théorie des courbes et des surfaces dans l'espace euclidien à trois dimensions, théorie qui est importante aussi bien par son contenu qu'en tant que source de représentations et d'exemples suggestifs pour la géométrie riemannienne et la géométrie de connexité affine. Mais pour la troisième année elle est d'une part trop élémentaire — à ce niveau les étudiants sont attirés par des concepts plus complexes — et d'autre part, pour remplir pleinement son rôle propédeudique, elle ne doit pas être délaissée trop rapidement au profit de la géométrie riemannienne.

Il est clair que l'exposé de cette théorie doit intervenir pas plus tard qu'au troisième semestre (et même au deuxième semestre comme je l'ai déjà proposé dans l'« Algèbre linéaire et géométrie différentielle » de ces « Leçons »). Par ailleurs, un cours de géométrie en troisième année n'est d'aucune utilité si les disciplines analytiques sont enseignées en deuxième année (et je suis convaincu que pour cette raison la géométrie finira par disparaître du programme de troisième année et peut-être même du programme universitaire).

Une solution radicale serait bien sûr dans la refonte du système traditionnel de répartition des cours. Mais vu que dans les conditions d'une lutte aiguë entre les chaires pour le nombre d'heures et les disciplines une telle refonte — qui doit être réalisée tôt ou tard — a peu de chances de succès, un palliatif serait de réintégrer la géométrie en deuxième année de façon à ce que les problèmes d'intégration soient clairement répartis entre l'analyse et la géométrie.

On pourrait proposer le schéma suivant. Après l'exposition de l'intégrale des fonctions sur des domaines de $\mathbb{R}^n$, le professeur d'analyse passerait le relais au professeur de géométrie pour développer l'intégration des densités et formes sur des variétés. Parallèlement, le professeur d'analyse illustrerait la théorie générale sur les cas particuliers des intégrales curvilignes et de surface de première (densité) et de deuxième (forme) espèce. En géométrie, on étudierait le théorème généralisé de Stokes qui se concrétise immédiatement en analyse par les formules de Green, Gauss-Ostrogradski et Stokes. Ce chassé-croisé conduirait à l'analyse vectorielle avec des éléments de théorie du potentiel, où le cours d'analyse se « jetterait » dans la théorie des intégrales impropres multidimensionnelles, et le cours de géométrie dans la théorie de la cohomologie. Ceci exige certes un degré d'orchestration qui n'est pas toujours facile à réaliser.

Comme les précédents ouvrages de cette série *) celui-ci a sa source dans les cours lus à « Mekhmat ». Il est consacré essentielle-

---

*) Voir M. Postnikov, *Leçons de géométrie:* Livre I, *Géométrie analytique*, Livre II, *Algèbre linéaire et géométrie différentielle*, Moscou, éd. Mir, 1981. (Dans le texte, on désigne ces livres par I et II. Le chiffre arabe qui suit I et II désigne le numéro de la leçon. (*N.d.R.*))

ment aux variétés différentiables, les notions de topologie générale ne sont pas mises en évidence mais insérées dans le texte.

Ces dernières années il s'est répandu sur les variétés différentiables un point de vue assez étrange qui, paradoxalement, est partagé par certains mathématiciens de renom. S'appuyant sur le fait qu'une variété différentiable peut être interprétée comme le résultat d'une tentative naturelle de généralisation axiomatique de la variété traitée comme un sous-ensemble de l'espace euclidien et définie par un système d'équations fonctionnellement indépendantes, ils arguent qu'étant donné qu'aux termes du théorème de plongement de Whitney, cette généralisation ne conduit pas à de nouveaux objets, il convient de définir la variété comme un tel sous-ensemble et que la notion générale de variété n'est qu'un exemple des multiples constructions axiomatiques qui apparaissent inévitablement lors de l'élaboration des notions mais qu'il vaut mieux ensuite oublier. Je ne peux en aucune façon souscrire à ce point de vue et j'estime qu'il est faux dans son principe ne serait-ce que parce qu'en pratique, par exemple en mécanique, les variétés apparaissent généralement sous une forme abstraite, ne sont plongées dans aucun espace euclidien et leur plongement de force induit — avec un grand degré d'arbitraire — une structure supplémentaire parfois utile, mais n'ayant en grande partie aucun rapport avec le problème étudié. Les adeptes de ce point de vue se réclament de Poincaré. En réalité, Poincaré a nettement saisi la nécessité d'un concept général de variété et s'est spécialement arrêté sur le collage des cartes d'un atlas. La référence à l'« extrémisme » de l'axiomatisation est un faux argument, car en réalité les variétés sont apparues non pas comme le résultat d'une « tentative naturelle de généralisation de la notion naïve de variété, définie par des équations » mais comme la réponse à la nécessité de définir rigoureusement un concept qui est nécessairement apparu dans le cadre de recherches mathématiques. La réalisation progressive de ces principes aurait rejeté les mathématiques en arrière, puisque par exemple vue sous cet angle l'algèbre linéaire n'aurait pas le droit d'exister sous sa forme actuelle en se fondant sur la notion d'espace vectoriel considéré comme le « résultat d'une tentative naturelle de généralisation de la notion naïve d'espace $\mathbb{R}^n$ » (ce qui est faux comme pour une variété), alors que le théorème d'isomorphisme montre que « cette généralisation ne conduit pratiquement pas à de nouveaux objets » (ce qui, bien que vrai, ne prive pas la notion d'espace vectoriel de sa valeur). C'est pourquoi dans cet ouvrage les variétés sont définies de façon habituelle, i.e. en s'appuyant sur la notion d'atlas, et les sous-ensembles des espaces euclidiens ne sont cités que comme exemples.

Les exercices proposés sont en règle générale triviaux et sont destinés exclusivement à un contrôle personnel du lecteur. Les plus difficiles sont donnés en petits caractères. Est également donné en

petits caractères un matériel auxiliaire se rapportant plus à l'algèbre ou l'analyse qu'à la géométrie.

Les cinq premières leçons qui ne concernent qu'indirectement la théorie des variétés différentiables sont consacrées à la géométrie différentielle élémentaire. Après la théorie des courbes (les formules de Frenet) on construit la première et la deuxième forme quadratique d'une surface, on établit les formules de dérivation de Weingarten et on prouve le théorème de Gauss d'invariance de la courbure totale. Tout ce qui ne conduit pas directement au théorème de Gauss a été oblitéré (les théorèmes de Meunier et d'Euler, les lignes géodésiques, les lignes asymptotiques, les lignes de courbure, etc.). En deuxième année ces questions sont parfois ajournées jusqu'au milieu du semestre (pour pouvoir développer les principales notions de théorie générale des variétés différentiables indispensables au cours d'équations différentielles). Ceci a certes permis d'éviter certaines répétitions (par exemple on n'a plus eu besoin de définir deux fois la différentielle d'une application différentiable: une fois pour les surfaces, une autre dans le cas général), mais s'est peu justifié méthodologiquement (et a trop subordonné la géométrie différentielle élémentaire — qui en principe est locale — à la théorie des variétés).

A proprement parler la théorie des variétés commence par la leçon 6. Les leçons 6 à 15 sont consacrées aux principales notions géométriques et aux théorèmes de la théorie des variétés. Si l'on veut limiter cet ouvrage à 11 leçons, on peut garder 4 des 5 premières leçons et sacrifier les leçons 8, 9 et 10 (qui développent essentiellement la théorie topologique de la dimension et le théorème de Tikhonov); dans un cours de 16 leçons on peut se séparer de la leçon 10. Les autres leçons de ce groupe (qui sont consacrées essentiellement aux théorèmes de Sard et de Whitney) doivent à mon avis être conservées dans toute variante.

Dans la variante à 11 leçons, le cours s'arrêterait là. Du reste il est possible dans ce cas, en abrégeant et comprimant l'exposé, d'économiser une heure et demie pour développer une partie des leçons 16 et 17. La théorie des formes différentielles (leçons 18, 19 et 20) serait reportée au prochain semestre (ou laissée à la discrétion du professeur d'analyse).

Dans la version à 16 leçons on peut éviter ce report et clôre le cours par la leçon 20 dans laquelle on démontre sur l'exemple de la sphère les diverses méthodes de calcul des groupes de cohomologie de de Rham. [A noter que cela reviendrait à exclure les leçons 24 à 29 qui seraient à la charge du professeur d'analyse.]

Dans les leçons 21, 22 et 23 on tente pour la première fois d'exposer d'une façon assez complète la théorie de l'homologie et de la cohomologie — jusqu'aux suites spectrales! On a pu le faire en rompant avec le point de vue traditionnel et en renonçant à l'exposé de la théorie simpliciale de l'homologie qui passe pour être géométrique-

ment suggestive. [Il m'est agréable de signaler que la même approche — mais à un niveau plus poussé — a été adoptée dans l'ouvrage de Bott R. & Tu L. « Differential forms in algebric topology » que nous recommandons vivement à ceux qui désireraient connaître les idées et les constructions fondamentales de la théorie classique de la cohomologie dans un esprit moderne.] Si l'on manque de temps, on peut sauter la deuxième moitié de la leçon 22 et la leçon 23.

Enfin les leçons 24 à 29 que l'on peut partiellement intervertir avec les leçons 21, 22 et 23 sont consacrées à l'intégration. L'exposé a été sciemment abrégé (on ne dit rien par exemple des fonctions d'ensembles additives), car ces leçons ne reflètent qu'une partie de la situation générale, l'autre partie se rapportant à l'analyse. La leçon 28 peut être entièrement confiée au professeur d'analyse. On peut aussi se limiter à la seule leçon 29 qui est indépendante des quatre leçons précédentes.

Mikhaïl Postnikov

LEÇON PREMIÈRE

Lignes simples du plan. — Définition d'une ligne par une équation. — Théorème de Whitney. — Courbes de Jordan. — Courbes régulières et différentiables. — Courbes non paramétrées. — Paramètre naturel.

Il existe plusieurs procédés de définition rigoureuse de la notion intuitive de ligne, procédés qui conduisent généralement à des résultats distincts. Néanmoins, dans les situations élémentaires toutes ces approches donnent pratiquement la même chose.

* * *

Etudions d'abord les lignes sur le plan.

Un ensemble Γ du plan s'appelle *graphique* s'il existe un système de coordonnées (euclidiennes ou affines) x, y et une fonction différentiable (ou continue) $f\colon I \to \mathbb{R}$ définie sur un intervalle (fermé, semi-ouvert ou ouvert) I de l'axe réel $\mathbb{R}$, telle que tout point de coordonnées x. y appartient à l'ensemble Γ si et seulement si $x \in I$ et $y = f(x)$. Intuitivement tous les graphiques sont manifestement des lignes.

On dit qu'un point p_0 d'un ensemble C du plan est *simple* s'il existe un disque ouvert U de centre p_0 tel que l'intersection $U \cap C$ soit un graphique.

Un ensemble C est *connexe* s'il ne peut être partagé en deux ensembles tels que tout point d'accumulation de l'un n'appartient pas à l'autre. (De façon imagée cela signifie que cet ensemble est composé d'un seul morceau.)

Un ensemble C du plan est une *ligne simple* s'il est connexe et composé uniquement de points simples.

Exercice 1. Montrer que tout graphique est connexe (donc, est une ligne simple).

Les procédés de définition rigoureuse de la notion de ligne diffèrent essentiellement par l'acception des points non simples. Nous escamoterons ces questions en convenant une fois pour toutes de n'envisager que des lignes simples.

Une ligne simple peut posséder (ou non) des *extrémités*. Elle ne peut avoir plus de deux extrémités. Une ligne simple à deux extrémités (de la forme d'un intervalle fermé courbé de l'axe numérique) est dite *fermée*, une ligne à une seule extrémité (de la forme d'un

intervalle semi-ouvert courbé de l'axe numérique), *semi-ouverte*. Une ligne simple sans extrémités peut avoir la forme soit d'un intervalle ouvert courbé de l'axe numérique, soit d'un cercle courbé. Dans le premier cas elle est dite *ouverte*, dans le second, *fermée*. (Par conséquent, le mot « fermé » est pris dans deux sens pour les lignes simples ! Cette ambiguïté historique ne doit pas être perdue de vue.)

* * *

Les lignes du plan sont définies le plus souvent par des équations de la forme

$$F(x, y) = 0, \tag{1}$$

où x, y sont des coordonnées (affines ou rectangulaires) du plan, et F une fonction de x, y. [Affirmer qu'un ensemble $\mathcal{L}$ est défini par une équation (1), signifie par définition qu'un point p du plan appartient à $\mathcal{L}$ si et seulement si ses coordonnées x, y satisfont à l'équation (1). Si F est traitée comme une fonction sur le plan, ceci signifie que $p \in \mathcal{L}$ si et seulement si $F(p) = 0$.]

Certes, pour obtenir une ligne (dans un sens ou dans l'autre) il faut soumettre la fonction F à certaines conditions. En premier lieu, il est naturel d'exiger qu'elle soit *continue*. [Si l'on envisage des fonctions discontinues, on peut à l'aide de l'équation (1) définir un ensemble arbitraire A de points du plan ; il suffit de prendre pour F la fonction $1 - \chi$, où χ est la *fonction caractéristique* de l'ensemble A, i.e. une fonction égale à 1 sur A et à 0 en dehors de A.]

Rappelons du cours d'analyse qu'un point p du plan (ou d'une façon générale d'un espace métrique, en particulier, euclidien) est un *point intérieur* d'un ensemble A s'il existe un $\varepsilon > 0$ tel que le ε-voisinage de ce point (une boule ouverte, ou un disque dans le plan, de rayon ε et de centre p) est entièrement contenu dans A. Un ensemble composé uniquement de points intérieurs est dit *ouvert*. Un ensemble C est *fermé* si son complémentaire est ouvert ou, ce qui est équivalent, si la limite de toute suite convergente de points $p_n \in C$ appartient aussi à C.

Un sous-ensemble d'un espace $\mathcal{A}$ affine (réel et de dimension finie) est *fermé* (resp. *ouvert*) s'il l'est pour une métrique euclidienne sur $\mathcal{A}$.

Exercice 2. Montrer que si un sous-ensemble d'un espace affine $\mathcal{A}$ est fermé (resp. ouvert) pour une métrique euclidienne sur $\mathcal{A}$, il l'est pour toute autre métrique.

Remarque 1. Une ligne simple fermée du plan est un ensemble fermé (et borné) dans chacune des deux acceptions du terme « ligne fermée ». En revanche, aucune ligne simple — même ouverte ! — n'est un ensemble ouvert. Plus, il existe des lignes ouvertes (le graphique de la tangente par exemple) qui sont des ensembles fermés (nécessairement non bornés).

Remarque 2. Il est immédiat de voir qu'*aucune ligne simple ne possède de points intérieurs.* Pour cette raison, Cantor a proposé de définir les lignes du plan comme des ensembles fermés sans points intérieurs. Cette définition est attrayante, mais elle est trop générale pour la plupart des théories mathématiques (elle occulte par exemple les lignes simples ouvertes).

Il est évident que si une fonction continue F est définie sur un espace métrique $\mathscr{X}$, l'ensemble de ses zéros est fermé. Ceci signifie qu'*une équation* (1) *avec une fonction F continue ne peut définir que des ensembles fermés du plan.*

Exercice 3. Montrer que, réciproquement, pour tout ensemble fermé C d'un espace métrique $\mathscr{X}$ il existe une fonction continue F définie sur $\mathscr{X}$ telle que $p \in C$ si et seulement si $F(p) = 0$. [N o t a. Envisager sur $\mathscr{X}$ la fonction

$$F(p) = \inf_{q \in C} \rho(p, q), \qquad p \in \mathscr{X},$$

qui donne la *distance de p à C*; ici ρ est la métrique de $\mathscr{X}$.]

Un point d'un ensemble (1) est dit *non singulier* ou encore *régulier* si en ce point les deux dérivées partielles

$$\frac{\partial F}{\partial x}, \quad \frac{\partial F}{\partial y}$$

de la fonction F existent, sont continues et non toutes deux nulles. Les autres points de l'ensemble (1) sont dits *singuliers.*

Le classique théorème des fonctions implicites affirme qu'au voisinage d'un point régulier tout ensemble (1) est un graphique, i.e. *tout point régulier est un point simple.* (La réciproque est fausse. Par exemple, pour $F(x, y) = x(x^2 + y^2)$ l'ensemble (1) est composé des points de l'axe des ordonnées $x = 0$, donc tous ses points sont simples. Pourtant le point (0, 0) est un point singulier.)

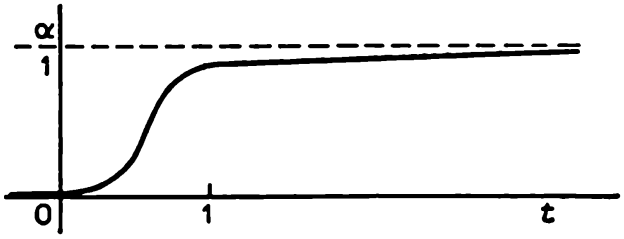
On en déduit que *l'ensemble de tous les points réguliers de tout ensemble* (1) *est la réunion de lignes simples* (se joignant en des points singuliers). Donc, si les points singuliers sont en nombre fini, un ensemble (1) correspond bien à l'idée intuitive d'une ligne. (Et il mérite bien son appellation.) Mais la situation change complètement s'il existe une grande quantité de points singuliers. Plus exactement, le mathématicien américain Whitney a prouvé que tout sous-ensemble fermé du plan peut être défini par une équation (1) avec une fonction F indéfiniment différentiable.

* * *

Le théorème de Whitney opère dans un espace affine $\mathscr{A}$ de points de dimension finie. Chaque fonction F peut être définie sur cet espace muni d'une origine O comme une fonction sur un espace vectoriel

associé $\mathcal{T}$, donc après choix d'une base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ dans $\mathcal{T}$, comme une fonction sur l'espace arithmétique $\mathbb{R}^n$. Une fonction F est une *fonction différentiable de classe* C^∞ si, traitée comme une fonction sur $\mathbb{R}^n$, elle admet des dérivées partielles continues de tout ordre. (Il est clair que si cette condition est remplie pour un repère $O\mathbf{e}_1 \ldots \mathbf{e}_n$, elle le sera pour tout autre repère.)

Théorème 1 (théorème Whitney). *Pour tout sous-ensemble fermé C d'un espace affine $\mathcal{A}$ il existe sur $\mathcal{A}$ une fonction F différentiable de classe C^∞ telle que $p \in C$ si et seulement si $F(p) = 0$.*



Graphique de la fonction α

La démonstration du théorème 1 s'appuie sur le lemme suivant qui présente un intérêt propre :

Lemme 1. *Il existe une fonction α: $\mathbb{R} \to \mathbb{R}$ monotone différentiable de classe C^∞ telle que*

1° $0 \leqslant \alpha(t) \leqslant 1$ *pour tout* $t \in \mathbb{R}$,

2° $\alpha(t) = 0$ *si et seulement si* $t \leqslant 0$.

D é m o n s t r a t i o n. Posons

$$\alpha(t) = \begin{cases} e^{-1/t} & \text{si } t > 0, \\ 0 & \text{sinon.} \end{cases} \tag{2}$$

Il est clair que la fonction (2) est monotone et possède les propriétés 1° et 2°. De plus il est évident que pour $t \neq 0$ elle est indéfiniment dérivable. Donc, il faut prouver seulement que cette fonction est indéfiniment dérivable pour $t = 0$.

Rappelons à cet effet qu'une fonction f définie au voisinage du point $t = 0$ est dérivable en ce point si existent les limites

$$\lim_{t \uparrow 0} \frac{f(t) - f(0)}{t}, \quad \lim_{t \downarrow 0} \frac{f(t) - f(0)}{t} \tag{3}$$

(les nombres dérivés à gauche et à droite de f en $t = 0$) et si ces limites sont égales.

D'autre part, si une fonction f est dérivable au voisinage de $t = 0$, sauf éventuellement en ce point, et si existent les limites

$$\lim_{t \uparrow 0} f'(t), \quad \lim_{t \downarrow 0} f'(t), \tag{4}$$

il résulte directement du théorème des accroissements finis de Lagrange que les limites (3) existent et sont égales à (4).

Puisque les limites à gauche (4) des fonctions $f = \alpha^{(n)}$ existent de toute évidence et sont nulles (car ces fonctions sont identiquement nulles pour $t < 0$), il suffit, pour démontrer le lemme 1, de montrer que pour tout $n \geqslant 0$ la limite

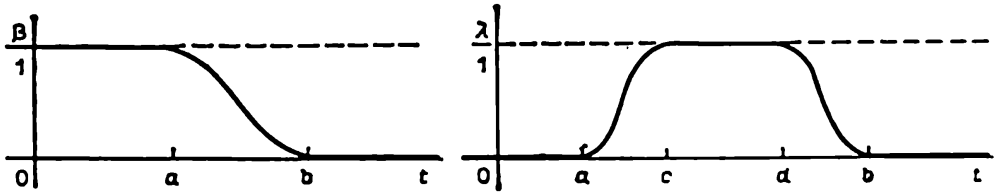
$$\lim_{t \downarrow 0} \alpha^{(n)}(t) = \lim_{t \downarrow 0} (e^{-1/t})^{(n)}$$

existe et est nulle.

Mais il est immédiat que pour tout $n \geqslant 0$ on a

$$(e^{-1/t})^{(n)} = e^{-1/t} p_n\left(\frac{1}{t}\right),$$

où $p_n = p_n(T)$ est un polynôme de degré $2n$. [Cette formule est



Graphique de la fonction β Graphique de la fonction λ

vraie pour $n = 0$; si elle l'est pour un $n \geqslant 0$, alors

$$(e^{-1/t})^{(n+1)} = \left[e^{-1/t} p_n\left(\frac{1}{t}\right)\right]' =$$
$$= e^{-1/t}\left[\frac{1}{t^2} p_n\left(\frac{1}{t}\right) - \frac{1}{t^2} p_n'\left(\frac{1}{t}\right)\right] = e^{-1/t} p_{n+1}\left(\frac{1}{t}\right),$$

où $p_{n+1}(T) = T^2 p_n(T) - T^2 p_n'(T)$ est un polynôme de degré $2n+$ $+2$].

Donc

$$\lim_{t \downarrow 0} (e^{-1/t})^{(n)} = \lim_{t \to +\infty} \frac{p_n(t)}{e^t} = 0,$$

c.q.f.d. □

Corollaire 1. *Pour tout segment* $[a, b]$ *de l'axe* $\mathbb{R}$ *il existe une fonction* $\beta : \mathbb{R} \to \mathbb{R}$ *différentiable de classe* C^∞ *telle que* $0 \leqslant \beta(t) \leqslant 1$ *pour tout* $t \in \mathbb{R}$ *et*

$$\beta(t) = \begin{cases} 1 & \text{pour } t \leqslant a, \\ 0 & \text{pour } t \geqslant b. \end{cases}$$

Démonstration. Il suffit de poser

$$\beta(t)=\frac{\alpha(b-t)}{\alpha(b-t)+\alpha(t-a)},$$

où α est la fonction du lemme 1. □

Remarque 3. Par analogie, on peut construire des fonctions de classe C^∞ ayant un comportement plus compliqué. Par exemple, pour tous nombres $a<c<d<b$ la formule

$$\lambda(t)=\frac{\alpha(B-|t-C|)}{\alpha(B-|t-C|)+\alpha(|t-D|-A)},$$

où

$$A=\frac{d-c}{2},\quad B=\frac{b-a}{2},\quad C=\frac{b+a}{2},\quad D=\frac{d+c}{2},$$

définit une fonction de classe C^∞, nulle à l'extérieur de $[a, b]$ et égale à l'unité sur $[c, d]$. Nous aurons besoin de cette fonction dans la leçon 15.

Soit $\mathscr{V}$ un espace vectoriel euclidien.

Notations. Pour tout $r>0$, par $\mathbb{B}_r^{\mathscr{V}}$ (ou simplement $\mathbb{B}_r$) on désignera une *boule* de rayon r et de centre 0 dans $\mathscr{V}$, i.e. l'ensemble de tous les vecteurs $\mathbf{x}\in\mathscr{V}$ tels que $|\mathbf{x}|\leqslant r$. La *boule ouverte* correspondante (l'ensemble des vecteurs $\mathbf{x}\in\mathscr{V}$ tels que $|\mathbf{x}|<r$) sera notée $\mathring{\mathbb{B}}_r^{\mathscr{V}}$.

Si $\mathscr{V}=\mathbb{R}^n$, on écrira généralement $\mathbb{B}_r^n$ au lieu de $\mathbb{B}_r^{\mathbb{R}^n}$.

Dans un espace ponctuel euclidien $\mathscr{A}$ la boule de rayon r et de centre p sera désignée par $\mathbb{B}_r^{\mathscr{A}}(p)$ (et la boule ouverte par $\mathring{\mathbb{B}}_r^{\mathscr{A}}(p)$).

Ces notations seront utilisées tout au long de cet ouvrage.

Corollaire 2. *Pour tout point p_0 d'un espace ponctuel euclidien $\mathscr{A}$ et tout $r>0$, il existe une fonction $f\colon \mathscr{A}\to\mathbb{R}$ telle que $0\leqslant f\leqslant 1$ sur $\mathscr{A}$ et*

$$f(p)=\begin{cases}1 \text{ si et seulement si } p\in\mathbb{B}_r(p_0),\\ 0 \text{ si et seulement si } p\notin\mathring{\mathbb{B}}_{2r}(p_0).\end{cases}$$

Démonstration. Il suffit de poser

$$f(p)=\beta(|\mathbf{x}|),$$

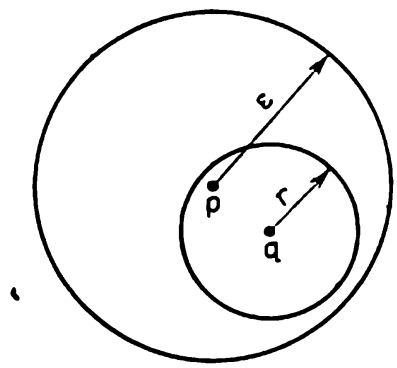
où $\mathbf{x}=\overrightarrow{p_0p}$ est le rayon vecteur du point p, d'origine p_0, et β la fonction du corollaire 1 construite pour le segment $[r, 2r]$. □

Rapportons $\mathscr{A}$ à un système de coordonnées rectangulaires. On dira qu'un point $p\in\mathscr{A}$ est *rationnel* si ses coordonnées le sont toutes et qu'une boule $\mathring{\mathbb{B}}_r(p)$ est *rationnelle* si son centre p et son rayon r le sont.

Lemme 2. *Tout ensemble ouvert $U \subset \mathcal{A}$ est la réunion d'une famille dénombrable (ou finie) de boules rationnelles, i.e. il existe des points rationnels $q_1, \ldots, q_m, \ldots$ et des nombres rationnels $r_1, \ldots, r_m, \ldots$ tels que*

$$U = \bigcup_{m=1}^{\infty} \mathring{\mathbb{B}}_{r_m}(q_m). \tag{5}$$

Démonstration. Par hypothèse, pour tout point $p \in U$ il existe un $\varepsilon > 0$ tel que $\mathring{\mathbb{B}}_\varepsilon(p) \subset U$. Considérons la boule rationnelle $\mathring{\mathbb{B}}_r(q)$, où q est un point rationnel tel que $\rho(p, q) < \varepsilon/2$, et r un nombre rationnel tel que $\rho(p, q) < r < \varepsilon/2$ (le point q et le



nombre r existent, car tout nombre réel peut être approché d'aussi près que l'on veut par des nombres rationnels). Puisque $\rho(p, q) < r$, on a $p \in \mathring{\mathbb{B}}_r(q)$, et comme

$$\rho(x, p) \leqslant \rho(x, q) + \rho(p, q) < 2r < \varepsilon$$

pour tout $x \in \mathring{\mathbb{B}}_r(q)$, il vient $\mathring{\mathbb{B}}_r(q) \subset \mathring{\mathbb{B}}_\varepsilon(p) \subset U$.

On voit donc que tout point $p \in U$ est contenu dans une boule rationnelle $\mathring{\mathbb{B}}_r(q) \subset U$. Ceci signifie que l'ensemble U est la réunion des boules rationnelles $\mathring{\mathbb{B}}_r(q)$ construites pour tous les points $p \in U$. Or l'ensemble des boules rationnelles de l'espace $\mathcal{A}$ est visiblement dénombrable. Donc l'ensemble des boules $\mathring{\mathbb{B}}_r(q)$ est au plus dénombrable. En les désignant par $\mathring{\mathbb{B}}_{r_m}(q_m)$ on obtient le développement (5). □

Tout est prêt maintenant pour la démonstration du théorème 1.

Démonstration du théorème 1. Sans perdre en généralité on peut de toute évidence admettre que l'espace affine

envisagé $\mathcal{A}$ est euclidien, donc le complémentaire $U = \mathcal{A} \setminus C$ de l'ensemble C se représente sous la forme (5). Soit f_m la fonction du corollaire 2 du lemme 1 correspondant au point q_m et au nombre $r_m/2$. Cette fonction (donc chacune de ses dérivées partielles) est identiquement nulle à l'extérieur de l'ensemble compact (fermé et borné) $\mathbb{B}_{r_m}(q_m)$. Donc, pour tout $k \geqslant 0$ il existe un nombre $c_m^k > 0$ tel que le module de chaque dérivée partielle d'ordre k de la fonction f_m soit inférieur à c_m^k sur l'espace $\mathcal{A}$ tout entier. Soit

$$c_m = \max(1, c_m^0, c_m^1, \ldots, c_m^m).$$

Considérons la série de fonctions

$$\sum_{m=1}^{\infty} \frac{f_m}{2^m c_m}. \tag{6}$$

Puisque par construction $c_m \geqslant 1$ et $f_m \leqslant 1$, cette série est majorée par la série numérique convergente

$$\sum_{m=1}^{\infty} \frac{1}{2^m}. \tag{7}$$

Donc, la série (6) converge vers une fonction $F: \mathcal{A} \to \mathbb{R}$. Si $p \in C$, alors $p \notin \mathbb{B}_{r_m}(q_m)$ pour tout $m \geqslant 1$, et par conséquent $f_m(p) = 0$. Si $p \notin C$ (i.e. $p \in U$), il existe un $m \geqslant 1$ tel que $p \in \mathbb{B}_{r_m}(q_m)$. Alors $f_m(p) \neq 0$ et par suite $F(p) \neq 0$. Donc, $F(p) = 0$ si et seulement si $p \in C$.

D'autre part, puisque pour tout $m \geqslant k$ chaque dérivée partielle d'ordre k du m-ième terme de la série (6) est visiblement inférieure au m-ième terme de la série (7), alors en dérivant k fois la série (6), on obtient une série dont tous les termes, à l'exception possible des k premiers, sont majorés par ceux de la série (7) et cette série est donc uniformément convergente. En vertu du classique théorème de dérivation des séries, la somme de la série (6) est indéfiniment dérivable (et chacune de ses dérivées partielles est la somme de la série formée avec les dérivées partielles correspondantes de la série (6)).

Ce qui achève la démonstration du théorème 1. □

* * *

Une autre approche de la notion de ligne, liée généralement au mathématicien français Jordan, définit cette dernière comme la trajectoire d'un point mobile. Les lignes au sens de Jordan seront appelées courbes.

Suivant Jordan, on appelle *courbe* dans un espace affine $\mathcal{A}$ à n dimensions une application continue

$$\gamma: I \to \mathcal{A}, \tag{8}$$

où I est un intervalle de l'axe $\mathbb{R}$ (ouvert, semi-ouvert ou fermé), i.e. si un repère est choisi dans $\mathcal{A}$, une fonction vectorielle continue

$$\mathbf{r} = \mathbf{r}(t), \quad t \in I, \tag{9}$$

à valeurs dans l'espace vectoriel associé $\mathcal{V}$.

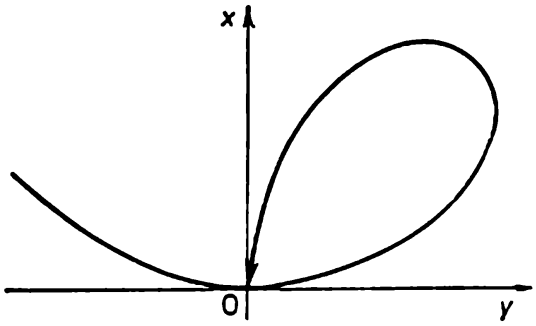
En coordonnées affines $x^1, \ldots, x^n$ la courbe de Jordan (8) est définie par les fonctions numériques continues

$$x^1 = x^1(t), \ldots, x^n = x^n(t), \quad t \in I. \tag{10}$$

Les équations (9) et (10) s'appellent *équations paramétriques* de la courbe (8) (resp. *vectorielle* et *analytiques*).

Signalons que contrairement aux lignes les courbes sont des applications et non des ensembles.

Mais dans la pratique il est plus commode de manipuler les courbes, du moins sur le plan de la terminologie, comme si elles



Folium de Descartes tronqué

étaient des ensembles. Par exemple pour tout $t_0 \in I$, le point $p_0 = \gamma(t_0)$ de l'espace $\mathcal{A}$ s'appelle *point de la courbe* (8) *associé à la valeur* t_0 *du paramètre* et on dit encore que *la courbe* (8) *passe par le point* p_0 *pour* $t = t_0$. Lorsque l'intervalle I est fermé (i.e. est de la forme $[a, b]$) les points $\gamma(a)$ et $\gamma(b)$ s'appellent *extrémités* de la courbe (9). On dit encore que la courbe (8) *relie* les points $\gamma(a)$ et $\gamma(b)$, etc.

Si $\gamma(a) = \gamma(b)$, la courbe (8) peut être traitée comme une application continue du cercle. De telles courbes sont dites *fermées*.

Si l'on veut faire une distinction entre une courbe et l'ensemble de ses points, on appelle ce dernier *support* de la courbe. Donc, le support de la courbe (8) n'est rien d'autre que l'image $\gamma(I)$ de l'intervalle I par l'application (8).

En général, le support d'une courbe peut se présenter sous une forme qui n'a rien de commun avec l'idée intuitive que l'on se fait d'une ligne. Il peut par exemple posséder des points intérieurs et même, comme le montre la célèbre courbe de Peano, remplir un carré.

Une courbe (8) est dite *simple* si, *primo*, c'est une application injective de I dans $\mathcal{A}$, i.e. $\gamma(t_1) = \gamma(t_2)$, $t_1, t_2 \in I$, si et seulement si $t_1 = t_2$, et, *secundo*, est une application bicontinue (on dit encore un *monéomorphisme*), i.e. telle que si pour une suite $\{t_m\}$ de points de l'intervalle I il existe un point $\tau \in I$ tel que $\lim \gamma(t_m) = \gamma(\tau)$, alors la suite $\{t_m\}$ est convergente (et $\lim t_m = \tau$). Une courbe fermée

$$\gamma : [a, b] \to \mathcal{A}, \quad \gamma(a) = \gamma(b), \tag{11}$$

est dite *simple* si $\gamma(t_1) = \gamma(t_2)$ pour $t_1 < t_2$ si et seulement si $t_1 = a$ et $t_2 = b$.

Un exemple typique d'application injective mais non monéomorphe d'un intervalle ouvert dans l'espace $\mathcal{A}$ est la courbe

$$x = \frac{3t^2}{1+t^3}, \quad y = \frac{3t}{1+t^3}, \quad -1 < t < +\infty$$

(« le folium de Descartes tronqué »).

Exercice 4. Montrer que toute application injective d'un intervalle $I = [a, b]$ dans $\mathcal{A}$ est un monéomorphisme.

Les supports des courbes simples s'appellent *arcs simples*.

Les arcs simples correspondent déjà à l'idée intuitive de ligne; en tout cas, du théorème d'invariance topologique de la dimension (cf. leçon 8) il résulte qu'ils ne possèdent pas de points intérieurs (pour $n > 1$ bien sûr). Pourtant leur structure peut être assez compliquée.

Exemple 1. Soient $x = x(t)$, $y = y(t)$, $0 \leqslant t \leqslant 1$, les équations paramétriques de la courbe de Peano dans le plan. Les équations

$$x = x(t), \quad y = y(t), \quad z = t, \quad 0 \leqslant t \leqslant 1,$$

définiront alors dans l'espace un arc simple dont la projection sur le plan Oxy est un carré. De façon imagée cela signifie que ce carré peut être recouvert par un toit qui n'est pas une surface mais une ligne !

* * *

On rappelle (cf. cours d'analyse) qu'une fonction réelle définie sur un intervalle $]a, b[$ est *différentiable de classe* C^r, où r est un entier naturel ou bien ∞, si elle possède des dérivées continues de tout ordre $\leqslant r$ (pour $r = \infty$ cette définition signifie qu'existent les dérivées partielles continues de tous les ordres; cf. plus haut). Pour cette raison, une courbe (8) définie sur un intervalle $I =]a, b[$ sera appelée *courbe différentiable de classe* C^r si toutes les fonctions (10) sont des fonctions de classe C^r. Puisque les dérivées

$$x^i(t)' = \frac{dx^i(t)}{dt}, \quad i = 1, \ldots, n,$$

des fonctions (10) sont les composantes du vecteur

$$\mathbf{r}'(t) = \lim_{h \to 0} \frac{\mathbf{r}(t+h) - \mathbf{r}(t)}{h}, \tag{12}$$

cette condition équivaut à l'existence, pour $r \neq \infty$, des dérivées continues de tout ordre $\leqslant r$ (pour $r = \infty$, des dérivées continues de tous les ordres) de la fonction vectorielle (9).

Dans la suite on admettra toujours que r est assez grand pour que les dérivations envisagées aient un sens et, d'une façon générale, on omettra de mentionner la classe C^r.

Lorsque l'intervalle I possède des extrémités (i.e. soit $I = [a, b]$, soit $I = [a, b[$, soit $I =]a, b]$), la courbe (8) est dite différentiable si elle est la restriction d'une courbe différentiable (de classe C^r) définie sur un intervalle $I' \supset I$.

Exercice 5. Montrer qu'il est équivalent de dire que (pour $r \neq \infty$) les fonctions (10) possèdent sur l'intervalle $]a, b[$ des dérivées continues de tout ordre $\leqslant r$ et aux points $t = a$ et/ou $t = b$ les dérivées unilatérales respectives.

Une courbe fermée (11) est dite différentiable si de plus les dérivées unilatérales en $t = a$ et $t = b$ sont confondues.

Le vecteur (12) s'appelle *vecteur tangent* à la courbe différentiable (11) au point t. En châtiant un peu le langage on l'appelle aussi vecteur tangent *au point* $\gamma(t)$. (Du reste cette terminologie est licite pour les courbes simples.)

A la leçon 15 nous prouverons le théorème de Sard qui entraîne en particulier que *le support d'une courbe différentiable ne possède pas de points intérieurs* (et même est un ensemble de mesure nulle). La projection d'une courbe différentiable étant de toute évidence une courbe différentiable, il s'ensuit que le phénomène décrit dans l'exemple 1 est impossible dans la classe des courbes différentiables.

Il est intéressant de noter qu'une courbe différentiable peut présenter des points anguleux.

Exemple 2. La courbe plane d'équations

$$x = \alpha(t), \quad y = \alpha(-t), \quad -\infty < t < +\infty, \tag{13}$$

où α est la fonction du lemme 1, possède un support composé de deux demi-droites de coordonnées $x = 0$, $y \geqslant 0$ et $x \geqslant 0$, $y = 0$ se joignant sous un angle droit !

Une courbe (8) (ou (11)) est dite *régulière en un point* t_0 si $\mathbf{r}'(t_0) \neq 0$. Une courbe régulière en tous ses points s'appelle *régulière*.

On remarquera que la courbe (13) n'est pas régulière au point anguleux $t = 0$, ce qui n'est pas un hasard, car on sait du cours d'analyse que *le support d'une courbe simple* (8) *régulière en un point* t_0 *possède une seule tangente au point* $\gamma(t_0)$ (de vecteur directeur $\mathbf{r}'(t_0)$).

* * *

Deux courbes

$$\gamma : I \to \mathcal{A}, \quad \gamma^* : I^* \to \mathcal{A}, \tag{14}$$

où I et I^* sont des intervalles de même type (tous deux fermés, ouverts ou semi-ouverts), sont dites *équivalentes* s'il existe une fonction dérivable (de classe C^r)

$$\varphi : I^* \to I \tag{15}$$

dont la dérivée est partout non nulle, et telle que $\gamma^* = \gamma \circ \varphi$, i.e. telle que

$$\gamma^* (t^*) = \gamma (\varphi (t^*)) \quad \text{pour tout} \quad t^* \in I^*. \tag{16}$$

On dit aussi que la fonction φ réalise un *changement de paramètre* sur la courbe γ.

Les classes d'équivalence des courbes s'appellent *courbes non paramétrées*. Pour distinguer les courbes des courbes non paramétrées on appellera parfois les premières *courbes paramétrées*.

Une courbe non paramétrée est dite *différentiable*, *simple* ou *régulière* selon qu'elle est la classe d'équivalence de courbes paramétrées différentiables, simples ou régulières. Cette définition est cohérente puisqu'une courbe équivalente à une courbe différentiable, simple ou régulière est visiblement différentiable ou respectivement simple et régulière.

Si des courbes (14) sont reliées par la relation (16), où φ est en général une fonction arbitraire, les supports de ces courbes sont confondus. Donc, *les courbes équivalentes possèdent le même support* (qui s'appelle support de la courbe non paramétrée correspondante), mais la réciproque est généralement fausse.

La situation est toutefois meilleure dans la classe des courbes simples et régulières.

Proposition 1. *Des courbes* (14) *simples et régulières possèdent le même support si et seulement si elles sont équivalentes.*

D é m o n s t r a t i o n. Si des courbes (14) sont simples et possèdent le même support, alors est définie de façon cohérente une application continue (pourquoi ?) $\varphi = \gamma^{-1} \circ \gamma^*$ de l'intervalle I^* sur l'intervalle I. Il nous faut donc prouver seulement que φ est dérivable et que sa dérivée est partout non nulle.

Soit t_0^* un point quelconque de l'intervalle I^* et soit $t_0 = \varphi (t_0^*)$. Si $p_0 = \gamma^* (t_0^*)$, alors $p_0 = \gamma (t_0)$ et le support des courbes (14) présente une seule tangente au point p_0. Ceci étant, si $\mathbf{r} = \mathbf{r} (t)$ et $\mathbf{r} = \mathbf{r}^* (t^*)$ sont les équations paramétriques vectorielles des courbes (14), les vecteurs $\mathbf{r}' (t_0)$ et $\mathbf{r}^{*\prime} (t_0^*)$ seront des vecteurs directeurs de cette tangente. Donc, *ces vecteurs sont colinéaires.*

La courbe γ étant régulière en t_0, on a $\mathbf{r}'(t_0) \neq 0$. Donc, si

$$x^1 = f^1(t), \ \ldots, \ x^n = f^n(t), \quad t \in I,$$

et

$$x^1 = g^1(t^*), \ \ldots, \ x^n = g^n(t^*), \quad t^* \in I^*,$$

sont les équations paramétriques des courbes (14), on peut, sans perdre en généralité, admettre que $\frac{df^1}{dt}(t_0) \neq 0$, donc que $\frac{dg^1}{dt}(t_0^*) \neq$ $\neq 0$, puisque les vecteurs $\mathbf{r}'(t_0)$ et $\mathbf{r}^{*\prime}(t_0^*)$ sont colinéaires.

Mais si $\frac{df^1}{dt}(t_0) \neq 0$, le classique théorème de la fonction réciproque nous dit que la fonction f^1 est localement inversible, i.e. il existe un intervalle $]a, b[$ de l'axe x contenant le point $x_0 = f^1(t_0)$ et sur cet intervalle une fonction $t = h(x)$ appliquant cet intervalle sur un intervalle $]\alpha, \beta[$ de l'axe t contenant le point t_0 (et contenu dans l'intervalle I) tels que

$$h(f^1(t)) = t \quad \text{pour tout point} \quad t \in]\alpha, \beta[.$$

Ceci étant, la fonction h est de même classe C^r que f^1, et sa dérivée n'est pas nulle en x_0.

Par construction, $g^1(t_0^*) = f^1(t_0) = x_0 \in]a, b[$. Il existe donc un intervalle $]\alpha^*, \beta^*[$ de l'axe t^* contenant le point t_0^* et contenu dans I^* tel que $g^1(t^*) \in]a, b[$ pour tout point $t^* \in]\alpha^*, \beta^*[$. Donc, sur l'intervalle $]\alpha^*, \beta^*[$ est définie la fonction

$$h \circ g^1 : t^* \longmapsto h(g^1(t^*)), \tag{17}$$

à valeurs dans l'intervalle $]\alpha, \beta[$. Cette fonction est de classe C^r et sa dérivée n'est pas nulle au point t_0^*.

D'un autre côté, par hypothèse,

$$f^i(\varphi(t^*)) = g^i(t^*)$$

pour tout point $t^* \in I^*$ et tout $i = 1, \ldots, n$; en particulier pour $t^* \in]\alpha^*, \beta^*[$ et $i = 1$. Donc, $\varphi(t^*) = (h \circ g^1)(t^*)$ pour $t^* \in$ $\in]\alpha^*, \beta^*[$, i.e. la fonction (17) est la restriction de φ à l'intervalle $]\alpha^*, \beta^*[$. Donc, la fonction φ est de classe C^r sur l'intervalle $]\alpha^*, \beta^*[$ et sa dérivée n'est pas nulle en t_0^*. Le point t_0^* étant un point arbitraire de l'intervalle I^* et les intervalles $]\alpha^*, \beta^*[$ recouvrant entièrement I^*, ceci prouve que la fonction φ est de classe C^r sur I^* et sa dérivée est non nulle sur I^*. □

La proposition 1 signifie que les courbes simples régulières sont à une équivalence près définies de façon unique par leurs supports (donc, peuvent leur être identifiées). Ces supports s'appellent *arcs simples réguliers*. Une courbe simple régulière, dont le support est un arc simple régulier $\mathcal{L}$, s'appelle aussi une *paramétrisation* de

l'arc $\mathcal{L}$. En principe, on identifiera les arcs simples réguliers à leurs paramétrisations (considérées à une équivalence près).

Remarque 4. Le problème, semble-t-il naturel, de l'extension de la proposition 1 à des courbes arbitraires moyennant des changements de paramètre plus généraux (par exemple, des changements à dérivées nulles) fait partie des (ô combien nombreux!) problèmes artificiels sans grand fondement.

Pour $n = 2$ (i.e. dans le plan) tout graphique est visiblement un arc simple régulier. Bien plus, le lecteur intéressé peut montrer que *dans le plan les arcs simples réguliers sont des lignes simples.* Donc, s'agissant des lignes simples, toutes les définitions rigoureuses de la notion intuitive de ligne aboutissent à un même résultat.

Exercice 6. Montrer que *dans le plan toute courbe régulière* (8) *est localement équivalente à un graphique*, i.e. pour tout point $t_0 \in I$ il existe dans $\mathbb{R}$ un voisinage $]a, b[\subset I$ tel que la courbe $\gamma \mid_{]a, b[}$ est équivalente à une courbe d'équations $x = t$, $y = f(t)$ (donc, est une courbe simple régulière).

Exercice 7. Citer un exemple montrant qu'une courbe régulière qui est une application injective peut néanmoins ne pas être une courbe simple (et peut même posséder un segment tout entier de points non simples).

* * *

Si l'espace $\mathcal{A}$ est euclidien, pour toute courbe différentiable (8) est définie sur I la fonction

$$t \mapsto |\mathbf{r}'(t)|, \quad t \in I,$$

qui représente la longueur du vecteur tangent $\mathbf{r}'(t)$. Cette fonction est visiblement continue, donc, lorsque $I = [a, b]$, il existe l'intégrale

$$s = \int_a^b |\mathbf{r}'(t)| \, dt.$$

On sait du cours d'analyse que cette intégrale est égale à la limite des longueurs des lignes polygonales inscrites dans la courbe (8), i.e. *à la longueur de la courbe* (8).

Supposons maintenant que l'intervalle I est arbitraire et soit $t_0 \in I$. La formule

$$s(t) = \int_{t_0}^{t} |\mathbf{r}'(t)| \, dt, \quad t \in I, \tag{18}$$

définit alors sur l'intervalle I une fonction différentiable appliquant I sur un intervalle J de l'axe s contenant 0. Cette fonction s'appelle *longueur d'arc.* (Signalons qu'elle peut prendre aussi des valeurs négatives.)

Si $s(t) = t - t_0$, le paramètre t est dit *naturel*. Donc, par abus de langage, on peut dire que *le paramètre t est naturel s'il est la longueur d'arc.*

La propriété du paramètre d'être naturel équivaut à l'identité $s'(t) = 1$. Puisque par définition $s'(t) = |\mathbf{r}'(t)|$, on voit que *le paramètre t sur la courbe* (8) *est naturel si et seulement si*

$$|\mathbf{r}'(t)| = 1 \quad \textit{pour tout} \quad t \in I.$$

On voit en particulier qu'*une courbe de paramètre naturel t est visiblement régulière.*

Réciproquement, supposons qu'une courbe (8) est régulière. Alors $|\mathbf{r}'(t)| > 0$ pour tout $t \in I$, et par conséquent, la fonction (19) est monotone et admet la fonction réciproque

$$\varphi: J \to I.$$

La courbe

$$\gamma_1 = \gamma \circ \varphi: J \to \mathcal{A}$$

est équivalente à la courbe γ et pour elle

$$\mathbf{r}_1'(s) = \mathbf{r}'(t)\frac{d\varphi}{ds}(s) = \mathbf{r}'(t)\frac{1}{s'(t)},$$

où $t = \varphi(s)$. Puisque $s'(t) = |\mathbf{r}'(t)|$, il s'ensuit que

$$|\mathbf{r}_1'(s)| = 1 \quad \text{pour tout} \quad s \in J,$$

i.e. le paramètre s est naturel sur la courbe γ_1.

On voit donc que *chaque courbe régulière est équivalente à une courbe de paramètre naturel.*

Vu qu'on s'est limité aux courbes régulières (et de plus simples), on peut donc, sans perdre en généralité, admettre que toutes les courbes envisagées sont de paramètre naturel. Il importe de noter que *pour un arc simple régulier le paramètre naturel est défini de façon unique à une transformation près de la forme*

$$t \mapsto \pm t + t_0$$

(i.e. au choix près de l'origine et de l'orientation).

Dans la suite, le paramètre naturel sera en principe désigné par le symbole s.

On désignera la dérivation par rapport à s par un point:

$$\dot{\mathbf{r}}(s) = \frac{d\mathbf{r}(s)}{ds}, \quad \ddot{\mathbf{r}}(s) = \frac{d^2\mathbf{r}(s)}{ds^2}, \quad \ldots, \text{ etc.}$$

On a déjà vu que la propriété pour un paramètre s d'être naturel équivaut à l'identité

$$|\dot{\mathbf{r}}(s)| = 1 \quad \text{pour tout } s.$$

De ce fait, il est utile de ne pas perdre de vue le lemme suivant (dans lequel s n'est certes pas un paramètre naturel) :

Lemme 3. *Soit* $\mathbf{u} = \mathbf{u}(s)$ *une fonction vectorielle dérivable telle que* $|\mathbf{u}(s)| = 1$ *pour tout* s. *Alors*

$$\mathbf{u}(s)\dot{\mathbf{u}}(s) = 0 \text{ pour tout } s. \tag{19}$$

Démonstration. L'égalité $|\mathbf{u}(s)| = 1$ équivaut à l'égalité $\mathbf{u}(s)^2 = 1$. Il est aisé de voir que le produit scalaire (de même d'ailleurs que le produit vectoriel) des vecteurs est justiciable de la formule ordinaire de dérivation d'un produit. En particulier

$$(\mathbf{u}^2)^{\cdot} = \dot{\mathbf{u}}\mathbf{u} + \mathbf{u}\dot{\mathbf{u}} = 2\mathbf{u}\dot{\mathbf{u}}.$$

Donc, si $\mathbf{u}^2 = 1$, alors $\mathbf{u}\dot{\mathbf{u}} = 0$. □

Corollaire. *Pour toute courbe* $\mathbf{r} = \mathbf{r}(s)$, *où* s *est le paramètre naturel, on a*

$$\dot{\mathbf{r}}(s)\ddot{\mathbf{r}}(s) = 0 \text{ pour tout } s. \tag{20}$$

LEÇON 2

Courbes dans le plan. — Formules de Frenet pour une courbe gauche. — Projections d'une courbe sur les plans de coordonnées du repère mobile. — Formules de Frenet pour une courbe dans un espace à n dimensions. — Existence et unicité d'une courbe de courbures données.

Soit

$$\gamma : I \to \mathcal{A} \tag{1}$$

une courbe simple et régulière dans un espace euclidien à n dimensions $\mathcal{A}$. On a vu à la leçon précédente qu'on pouvait, sans perdre en généralité, considérer que la courbe (1) est de paramètre naturel s.

Soit $\mathbf{r} = \mathbf{r}(s)$ l'équation paramétrique vectorielle de la courbe (1) et soit

$$\mathbf{t}(s) = \dot{\mathbf{r}}(s) \tag{2}$$

un vecteur tangent à cette courbe. Le paramètre s étant naturel, le vecteur (2) est un vecteur unitaire, et le vecteur

$$\dot{\mathbf{t}}(s) = \ddot{\mathbf{r}}(s)$$

lui est orthogonal :

$$\mathbf{t}(s)\,\dot{\mathbf{t}}(s) = 0 \text{ pour tout } s.$$

Définition 1. La longueur $|\dot{\mathbf{t}}(s)|$ du vecteur $\dot{\mathbf{t}}(s)$, qui est désignée par $k(s)$ (ou simplement par k), s'appelle *courbure* de la courbe (1) au point s (ou au point de rayon vecteur $\mathbf{r}(s)$).

Pour une courbe plane par exemple

$$k(s) = \sqrt{\ddot{x}^2(s) + \ddot{y}^2(s)},$$

où $x = x(s)$ et $y = y(s)$ sont les équations paramétriques analytiques de la courbe (1) dans un système de coordonnées euclidien x, y.

On appelle *courbure d'une courbe* de paramétrage arbitraire la courbure d'une courbe équivalente de paramétrage naturel. La formule de cette courbure (que l'on peut établir par des calculs simples mais assez volumineux à l'aide des seules formules de dérivation des fonctions) est assez compliquée même pour les courbes simples :

$$k = \left| \frac{x''y' - y''x'}{[(x')^2 + (y')^2]^{3/2}} \right|.$$

De façon imagée le nombre k est la vitesse instantanée de rotation du vecteur unitaire $\mathbf{t}$. Il est clair que cette vitesse est d'autant plus grande que la courbe est « courbée »; d'où le nom de courbure.

Si le plan est muni d'une orientation, on peut envisager la *courbure relative* k_{rel} qui est égale à k ($\neq 0$) ou à $-k$ selon que les vecteurs $\mathbf{t}$ et $\dot{\mathbf{t}}$ forment une base directe ou non. Nous aurons besoin de la courbure à la leçon 4.

Exemple 1. Si

$$x = x_0 + sl, \quad y = y_0 + sm, \quad \text{où} \quad l^2 + m^2 = 1,$$

i.e. si la courbe envisagée est une droite, alors $\ddot{x} = 0$ et $\ddot{y} = 0$. Donc, $k = 0$ pour tout s, i.e. comme on s'y attendait, *la courbure d'une droite est identiquement nulle.*

Puisque les fonctions linéaires (et affines) sont les seules fonctions dont la dérivée seconde soit identiquement nulle, la réciproque est également vraie, i.e. *toute courbe dont la courbure est identiquement nulle est une droite* (ou un segment de droite).

Le point $\mathbf{r}_0 = \mathbf{r}(s_0)$ de la courbe $\mathbf{r} = \mathbf{r}(s)$ tel que $k(s_0) = 0$ s'appelle *point méplat.*

Exemple 2. Les équations paramétriques d'un cercle de rayon R sont, de toute évidence, de la forme

$$x = R\cos\frac{s}{R}, \quad y = R\sin\frac{s}{R},$$

où s est un paramètre naturel. Comme

$$\ddot{x} = -\frac{1}{R}\cos\frac{s}{R}, \quad \ddot{y} = -\frac{1}{R}\sin\frac{s}{R},$$

on a

$$k(s) = \frac{1}{R} \quad \text{pour tout } s.$$

Donc, *la courbure du cercle est constante et égale à l'inverse de son rayon.*

Du théorème général 1 prouvé plus bas il s'ensuit que réciproquement *toute courbe de courbure constante est un cercle* (ou un arc de cercle).

Si $k(s_0) \neq 0$ pour une courbe, le nombre $R(s_0) = \frac{1}{k(s_0)}$ s'appelle *rayon de courbure* de cette courbe en s_0.

Une courbe $\mathbf{r} = \mathbf{r}(s)$ s'appelle *courbe générique* si elle ne présente aucun point méplat, i.e. si $k(s) \neq 0$ pour tout s. En tout point d'une telle courbe est défini le vecteur unitaire

$$\mathbf{n}(s) = \frac{\dot{\mathbf{t}}(s)}{k(s)},$$

dirigé suivant la *normale* à la courbe (i.e. suivant la droite perpendiculaire à la tangente au point de contact).

Les vecteurs $\mathbf{t}(s)$ et $\mathbf{n}(s)$ forment pour tout s une base orthonormée appelée *base mobile de Frenet*. Les mêmes vecteurs rapportés à un point de rayon vecteur $\mathbf{r}(s)$ constituent un *repère mobile de Frenet*.

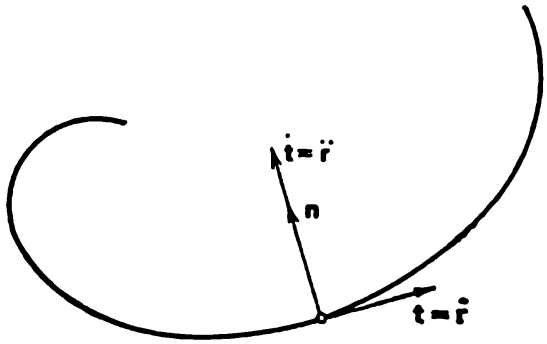
Par définition,

$$\dot{\mathbf{t}}(s) = k(s)\,\mathbf{n}(s).$$

Etablissons une formule analogue pour le vecteur $\dot{\mathbf{n}}(s)$. Soit

$$\dot{\mathbf{n}}(s) = \alpha(s)\,\mathbf{t}(s) + \beta(s)\,\mathbf{n}(s)$$

la décomposition de ce vecteur suivant la base $\mathbf{t} = \mathbf{t}(s)$, $\mathbf{n} = \mathbf{n}(s)$. Puisque $\mathbf{tn} = 0$, on a $\dot{\mathbf{t}}\mathbf{n} + \mathbf{t}\dot{\mathbf{n}} = 0$ (on s'est encore servi du fait



Repère mobile de Frenet d'une courbe plane

que le produit scalaire de vecteurs est justiciable de la formule ordinaire de dérivation d'un produit), donc $\alpha = \mathbf{t}\dot{\mathbf{n}} = -\dot{\mathbf{t}}\mathbf{n} = -k$. D'autre part, $\beta = \mathbf{n}\dot{\mathbf{n}} = 0$ en vertu du lemme 2 de la leçon 1. Ceci prouve que *pour toute courbe générique on a les formules*

$$(3)\qquad \begin{aligned} \dot{\mathbf{t}} &= k\mathbf{n}, \\ \dot{\mathbf{n}} &= -k\mathbf{t} \end{aligned}$$

(on omet d'écrire l'argument s) qui décrivent la rotation instantanée de la base mobile de Frenet.

Les formules (3) s'appellent *formules de Frenet pour une courbe plane*.

Remarque 1. Dans un plan orienté, on peut définir un repère de Frenet pour les courbes sans points méplats en prenant pour $\mathbf{n}(s)$ le vecteur qui forme avec $\mathbf{t}(s)$ une base directe. Dans les formules (3), il faudra alors remplacer la courbure k par la courbure relative k_{rel}.

* * *

Dans un espace rapporté à un système de coordonnées rectangulaires x, y, z, la formule de courbure s'écrit

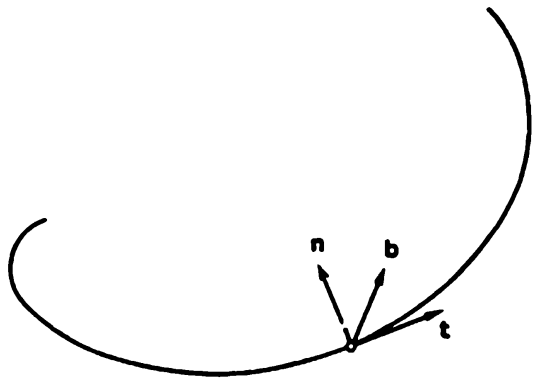
$$k = \sqrt{\ddot{x}^2 + \ddot{y}^2 + \ddot{z}^2}.$$

Dans l'espace comme dans le plan, la courbe (1) s'appelle *courbe générique* si $k(s) \neq 0$ pour tout s. Pour une telle courbe est défini le vecteur unitaire

$$\mathbf{n}(s) = \frac{\dot{\mathbf{t}}(s)}{k(s)},$$

appelé *vecteur unitaire de la normale principale* à la courbe.

Nous pouvons maintenant (dans l'hypothèse que l'espace est orienté) introduire un troisième vecteur $\mathbf{b}(s)$ formant avec les vecteurs $\mathbf{t}(s)$ et $\mathbf{n}(s)$ une base directe orthonormée $\mathbf{t}(s)$, $\mathbf{n}(s)$, $\mathbf{b}(s)$ (i.e. une base telle que $\mathbf{b}(s) = \mathbf{t}(s) \times \mathbf{n}(s)$). Ce vecteur s'appelle *vecteur binormal* et la base $\mathbf{t}(s)$, $\mathbf{n}(s)$, $\mathbf{b}(s)$, *base mobile de Frenet* attachée à cette courbe générique. Les vecteurs $\mathbf{t}(s)$, $\mathbf{n}(s)$, $\mathbf{b}(s)$ rapportés à un point de rayon vecteur $\mathbf{r}(s)$ constituent un *repère mobile de Frenet*.



Repère mobile de Frenet d'une courbe gauche

Par construction (on omet l'argument s pour simplifier les formules)

$$\dot{\mathbf{t}} = k\mathbf{n}.$$

De plus, puisque $\mathbf{b} = \mathbf{t} \times \mathbf{n}$, il vient

$$\dot{\mathbf{b}} = \dot{\mathbf{t}} \times \mathbf{n} + \mathbf{t} \times \dot{\mathbf{n}} = \mathbf{t} \times \dot{\mathbf{n}},$$

d'où il s'ensuit que $\dot{\mathbf{b}}\mathbf{t} = 0$. Ce qui prouve, puisque $\dot{\mathbf{b}}\mathbf{b} = 0$ d'après le lemme 2 de la leçon 1, que le vecteur $\dot{\mathbf{b}}$ est colinéaire au vecteur $\mathbf{n}$, i.e. il existe un nombre $\varkappa = \varkappa(s)$ tel que

$$\dot{\mathbf{b}} = -\varkappa\mathbf{n}.$$

Le nombre $\varkappa(s)$ s'appelle *torsion* de la courbe au point $\mathbf{r}(s)$. C'est la vitesse de rotation du vecteur binormal.

Une dérivation des égalités $\mathbf{nt} = 0$ et $\mathbf{nb} = 0$ nous donne $\dot{\mathbf{n}}\mathbf{t} = -\mathbf{n}\dot{\mathbf{t}} = -k$ et $\dot{\mathbf{n}}\mathbf{b} = -\mathbf{n}\dot{\mathbf{b}} = \varkappa$. Comme $\dot{\mathbf{n}}\mathbf{n} = 0$ (lemme 2 de la leçon 1), on a ainsi prouvé que

$$\dot{\mathbf{n}} = -k\mathbf{t} + \varkappa\mathbf{b}.$$

Donc, *pour toute courbe générique on a les formules*

$$\begin{aligned} \dot{\mathbf{t}} &= k\mathbf{n}, \\ \dot{\mathbf{n}} &= -k\mathbf{t} + \varkappa\mathbf{b}, \\ \dot{\mathbf{b}} &= -\varkappa\mathbf{n}, \end{aligned} \tag{4}$$

appelées *formules de Frenet pour une courbe gauche.*

Exemple 3. Si la courbe $\mathbf{r} = \mathbf{r}(s)$ est située dans un plan Π, les vecteurs $\dot{\mathbf{r}}(s)$ et $\ddot{\mathbf{r}}(s)$ sont parallèles à Π (car il en est ainsi pour les accroissements $\mathbf{r}(s + \Delta s) - \mathbf{r}(s)$ et $\dot{\mathbf{r}}(s + \Delta s) - \dot{\mathbf{r}}(s)$ des vecteurs $\mathbf{r}(s)$ et $\dot{\mathbf{r}}(s)$). Donc, les vecteurs $\mathbf{t}(s)$ et $\mathbf{n}(s)$ sont parallèles à Π et par suite $\mathbf{b}(s) \perp \Pi$. Ceci prouve que $\mathbf{b}(s) = \text{const}$, donc que $\varkappa(s) = 0$ pour tout s. Réciproquement, si $\varkappa(s) = 0$ pour tout s, c'est que $\mathbf{b}(s) = \mathbf{b}_0 = \text{const}$. Alors $(\mathbf{r}(s)\,\mathbf{b}_0)^{\cdot} = \dot{\mathbf{r}}(s)\,\mathbf{b}_0 = \mathbf{t}(s)\,\mathbf{b}_0 = 0$ pour tout s, donc $\mathbf{r}(s)\,\mathbf{b}_0 = \text{const}$. Ceci exprime que la courbe $\mathbf{r} = \mathbf{r}(s)$ est située dans le plan $\mathbf{r}\mathbf{b}_0 = \text{const}$. Donc, *une courbe de l'espace est une courbe plane si et seulement si sa torsion est identiquement nulle.*

Exemple 4. L'*hélice* est la trajectoire d'un point se déplaçant à une vitesse constante le long d'une génératrice d'un cylindre circulaire droit animé d'une rotation uniforme autour de son axe. Les équations paramétriques de cette courbe sont

$$x = a\cos t, \quad y = a\sin t, \quad z = bt.$$

Puisque $x' = -a\sin t$, $y' = a\cos t$, $z' = b$, on a

$$s' = \sqrt{(x')^2 + (y')^2 + (z')^2} = \sqrt{a^2 + b^2},$$

donc $s = ct$, où $c = \sqrt{a^2 + b^2}$. Par suite

$$x = a\cos\frac{s}{c}, \quad y = a\sin\frac{s}{c}, \quad z = \frac{b}{c}s.$$

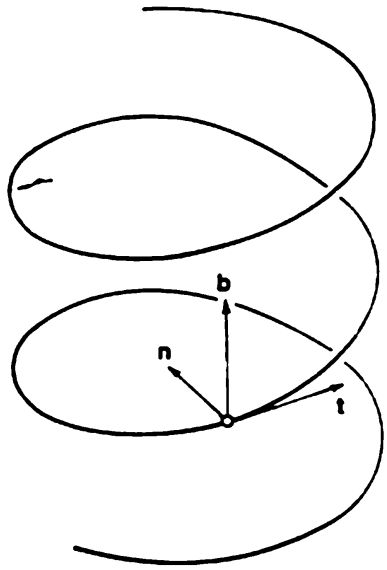
Mais alors

$$\dot{x} = -\frac{a}{c}\sin\frac{s}{c}, \quad \dot{y} = \frac{a}{c}\cos\frac{s}{c}, \quad \dot{z} = \frac{b}{c},$$

$$\ddot{x} = -\frac{a}{c^2}\cos\frac{s}{c}, \quad \ddot{y} = -\frac{a}{c^2}\sin\frac{s}{c}, \quad \dot{z} = 0,$$

donc

$$k = \sqrt{\ddot{x}^2 + \ddot{y}^2 + \ddot{z}^2} = \frac{a}{c^2} = \text{const.}$$



Hélice

Par ailleurs,

$$\mathbf{t} = \left(-\frac{a}{c}\sin\frac{s}{c}, \ \frac{a}{c}\cos\frac{s}{c}, \ \frac{b}{c}\right),$$

$$\mathbf{n} = \left(-\cos\frac{s}{c}, \ -\sin\frac{s}{c}, \ 0\right)$$

et

$$\mathbf{b} = \mathbf{t} \times \mathbf{n} = \left(\frac{b}{c}\sin\frac{s}{c}, \ -\frac{b}{c}\cos\frac{s}{c}, \ \frac{a}{c_z}\right).$$

Donc

$$\dot{\mathbf{b}} = \left(\frac{b}{c^2}\cos\frac{s}{c}, \ \frac{b}{c^2}\sin\frac{s}{c}, \ 0\right) = -\frac{b}{c^2}\mathbf{n},$$

et par suite

$$\varkappa = \frac{b}{c^2} = \text{const.}$$

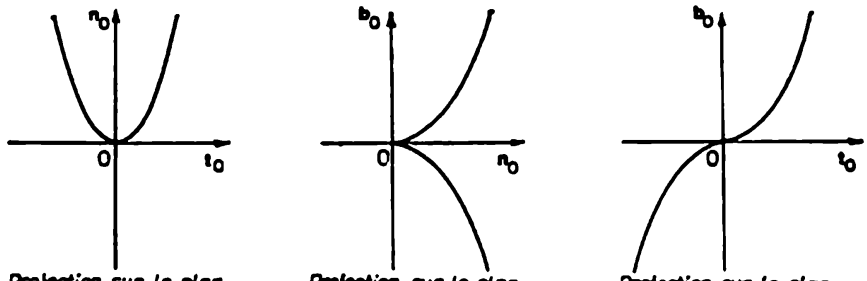
Par conséquent, *la courbure et la torsion de l'hélice sont constantes.*

En vertu du théorème général 1 prouvé plus bas on a la réciproque: *toute courbe dont la courbure et la torsion sont constantes est une hélice* (ou un arc d'hélice).

Remarque 2. A noter que la notion de courbe générique n'admet pas la même interprétation dans l'espace et dans le plan. Pour réaliser cette unité, il faut, dans le cas des courbes planes, envisager la courbure relative et non pas la courbure absolue. Cf. remarque 1.

* * *

Pour étudier une courbe gauche $\mathbf{r} = \mathbf{r}(s)$ au voisinage de l'un quelconque de ses points, plaçons l'origine O des coordonnées en ce



Projection sur le plan osculateur — Projection sur le plan normal — Projection sur le plan rectifiant

point et prenons pour base $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ la base mobile $\mathbf{t}_0$, $\mathbf{n}_0$, $\mathbf{b}_0$ au point O. On a alors

$$\mathbf{r}(0) = 0, \quad \dot{\mathbf{r}}(0) = \mathbf{t}_0 = \mathbf{i}, \quad \ddot{\mathbf{r}}(0) = k_0\mathbf{n}_0 = k_0\mathbf{j},$$

$$\dddot{\mathbf{r}}(0) = \dot{k}_0\mathbf{n}_0 + k_0\dot{\mathbf{n}}_0 = -k_0^2\mathbf{i} + \dot{k}_0\mathbf{j} + k_0\varkappa_0\mathbf{k},$$

où k_0, $\dot{k}_0$ et $\varkappa_0$ sont les valeurs des fonctions k, $\dot{k}$ et $\varkappa$ pour $s = 0$. La formule de Taylor nous donne donc

$$\mathbf{r}(s) = \mathbf{r}(0) + s\dot{\mathbf{r}}(0) + \frac{s^2}{2}\ddot{\mathbf{r}}(0) + \frac{s^3}{6}\dddot{\mathbf{r}}(0) + \ldots =$$

$$= (s + \ldots)\,\mathbf{i} + \left(\frac{k_0}{2}s + \ldots\right)\mathbf{j} + \left(\frac{k_0\varkappa_0}{6}s^3 + \ldots\right)\mathbf{k},$$

ce qui exprime qu'au voisinage du point O la courbe $\mathbf{r} = \mathbf{r}(s)$ est définie par les équations paramétriques

$$x = s + \ldots$$

$$y = \frac{k_0}{2}s^2 + \ldots$$

$$z = \frac{k_0\varkappa_0}{6}s^3 + \ldots$$

Si $k_0 \neq 0$, $\varkappa_0 \neq 0$, la projection de la courbe sur le plan $O\mathbf{ij} = O\mathbf{t}_0\mathbf{n}_0$ (appelé *plan osculateur* de la courbe au point O) est approximativement confondue avec la parabole

$$x = s, \quad y = \frac{k_0}{2} s^2;$$

la projection sur le plan $O\mathbf{jk} = O\mathbf{n}_0\mathbf{b}_0$ (appelé *plan normal* de la courbe au point O), avec la parabole semi-cubique

$$y = \frac{k_0}{2} s^2, \quad z = \frac{k_0\varkappa_0}{6} s^3;$$

et enfin la projection sur le plan $O\mathbf{ik} = O\mathbf{t}_0\mathbf{b}_0$ (appelé *plan rectifiant* de la courbe au point O), avec la parabole cubique

$$x = s, \quad z = \frac{k_0\varkappa_0}{2} s^3.$$

Ceci nous donne une idée assez suggestive sur la structure d'une courbe gauche au voisinage de l'un quelconque de ses points (en lesquels sont non nulles la courbure et la torsion).

* * *

Considérons maintenant le cas général d'un espace euclidien de dimension quelconque $n \geqslant 2$.

Une courbe $\mathbf{r} = \mathbf{r}(s)$ de paramétrage s dans un espace euclidien orienté à n dimensions est dite *générique* si pour tout s les vecteurs

$$\dot{\mathbf{r}}(s), \ldots, \overset{(n-1)}{\mathbf{r}}(s) \tag{5}$$

sont linéairement indépendants.

En appliquant aux vecteurs (5) le procédé d'orthogonalisation de Gram-Schmidt, on obtient un système orthonormal de vecteurs $\mathbf{t}_1(s), \ldots, \mathbf{t}_{n-1}(s)$. Soit $\mathbf{t}_n(s)$ le vecteur (unique) tel que

$$\mathbf{t}_1(s), \ldots, \mathbf{t}_{n-1}(s), \mathbf{t}_n(s) \tag{6}$$

soit une base orthonormale directe.

Définition 2. La base (6) s'appelle *base mobile de Frenet* de la courbe générique.

Soit

$$\dot{\mathbf{t}}_i = \sum_{j=1}^{n} \alpha_{ij}\mathbf{t}_j, \quad i = 1, \ldots, n$$

(pour simplifier l'écriture on omet l'argument s). Puisque par construction le vecteur $\mathbf{t}_i$, $i = 1, \ldots, n-1$, s'exprime linéairement en fonction des vecteurs $\dot{\mathbf{r}}, \ldots, \overset{(i)}{\mathbf{r}}$, le vecteur $\dot{\mathbf{t}}_i$ s'exprime linéaire-

ment en fonction des vecteurs $\dot{\mathbf{r}}, \ldots, \overset{(i+1)}{\mathbf{r}}$. Les derniers vecteurs étant des combinaisons linéaires des vecteurs $\mathbf{t}_1, \ldots, \mathbf{t}_{i+1}$, ceci prouve que $\alpha_{ij} = 0$ pour $j > i + 1$.

D'autre part, puisque $\mathbf{t}_i\mathbf{t}_j = \delta_{ij}$, on a $\dot{\mathbf{t}}_i\mathbf{t}_j + \mathbf{t}_i\dot{\mathbf{t}}_j = 0$, i.e.

$$\alpha_{ij} + \alpha_{ji} = 0.$$

Donc, $\alpha_{ii} = 0$ et $\alpha_{ij} = 0$ pour $j < i - 1$.

Par conséquent, seuls les coefficients $\alpha_{i,\, i+1} = -\alpha_{i+1,\, i}$ peuvent être non nuls. En posant

$$k_1 = \alpha_{12}, \quad k_2 = \alpha_{23}, \ \ldots, \ k_{n-1} = \alpha_{n-1,\, n},$$

on voit que

$$(7) \qquad \begin{aligned} \dot{\mathbf{t}}_1 &= k_1\mathbf{t}_2, \\ \dot{\mathbf{t}}_2 &= -k_1\mathbf{t}_1 + k_2\mathbf{t}_3, \\ &\ldots\ldots\ldots\ldots \\ \dot{\mathbf{t}}_{n-1} &= -k_{n-2}\mathbf{t}_{n-2} + k_{n-1}\mathbf{t}_n, \\ \dot{\mathbf{t}}_n &= -k_{n-1}\mathbf{t}_{n-1}. \end{aligned}$$

Ces formules s'appellent *formules de Frenet pour une courbe dans un espace à n dimensions.*

Les fonctions $k_1 = k_1(s), \ldots, k_{n-1} = k_{n-1}(s)$ s'appellent *courbures* de cette courbe. Signalons qu'elles ne sont définies que pour les courbes génériques.

Dans les formules

$$(8) \qquad \mathbf{t}_i = \beta_{i1}\dot{\mathbf{r}} + \ldots + \beta_{ii}\overset{(i)}{\mathbf{r}}_i, \quad i = 1, \ldots, n-1,$$

acquises par le procédé d'orthogonalisation de Gram-Schmidt, les derniers coefficients β_{ii} sont strictement positifs. Donc, il en est de même des coefficients $\gamma_{ii} = \beta_{ii}^{-1}$ dans les formules inverses

$$(9) \qquad \overset{(i)}{\mathbf{r}} = \gamma_{i1}\mathbf{t}_1 + \ldots + \gamma_{ii}\mathbf{t}_i.$$

Une dérivation de la formule (8) nous donne

$$\dot{\mathbf{t}}_i = \dot{\beta}_{i1}\dot{\mathbf{r}} + (\dot{\beta}_{i2} + \beta_{i1})\ddot{\mathbf{r}} + \ldots + (\dot{\beta}_{ii} + \beta_{i,\, i-1})\overset{(i)}{\mathbf{r}} + \beta_{ii}\overset{(i+1)}{\mathbf{r}},$$
$$i = 1, \ldots, n-1.$$

En y remplaçant (pour $i < n - 1$) les vecteurs $\dot{\mathbf{r}}, \ldots, \overset{(i+1)}{\mathbf{r}}$ par leurs expressions (9), on obtient les formules (7), ce qui montre que

$$k_i = \beta_{ii}\gamma_{i+i,\, i+1}, \quad i = 1, \ldots, n-2.$$

Il s'ensuit en particulier que *les courbures $k_1, \ldots, k_{n-2}$ de toute courbe générique sont strictement positives.* La courbure k_{n-1} peut être de signe quelconque.

* * *

Montrons maintenant que toutes $n-1$ fonctions

$$k_1(s) > 0, \ \ldots, \ k_{n-2}(s) > 0, \quad k_{n-1}(s) \tag{10}$$

(définies sur un intervalle $]a, b[$ de l'axe $\mathbb{R}$) peuvent servir de courbures d'une courbe (régulière mais généralement non simple) et que ces courbures définissent cette courbe de façon unique (à une congruence près).

Supposons pour fixer les idées que $a < 0 < b$.

Théorème 1. *Sur l'intervalle* $]a, b[$ *soient données* n *fonctions différentiables* (10) *toutes strictement positives à l'exception possible de la dernière. Pour toute origine* $O \in \mathcal{A}$ *et toute base orthonormée directe* $\mathbf{i}_1, \ldots, \mathbf{i}_n$ *il existe une courbe générique* $\mathbf{r} = \mathbf{r}(s)$, $a < s < b$, *et une seule, jouissant des deux propriétés suivantes*:

1° *les courbures de cette courbe sont les fonctions* (10);

2° *pour* $s = 0$ *on a*

$$\mathbf{r}(0) = 0, \quad \mathbf{t}_1(0) = \mathbf{i}_1, \ \ldots, \ \mathbf{t}_n(0) = \mathbf{i}_n.$$

Démonstration. Nous la ferons en quatre étapes.

Etape 1. On se servira du classique théorème d'existence et d'unicité de la solution (EUS) des équations différentielles linéaires, qui se prouve en théorie des équations différentielles.

Théorème d'EUS. *Sur un intervalle* $]a, b[$ *soient données* m^2 *fonctions différentiables* $A_{ij}(s)$, $i, j = 1, \ldots, m$, *et des nombres arbitraires* $x_1^{(0)}, \ldots, x_m^{(0)}$. *Il existe alors une famille de fonctions différentiables* $x_1(s), \ldots, x_m(s)$, $a < s < b$, *et une seule, telle que*

1° *les relations*

$$\begin{aligned} \dot{x}_1 &= A_{11}x_1 + \ldots + A_{1m}x_m, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ \dot{x}_m &= A_{m1}x_1 + \ldots + A_{mm}x_m \end{aligned} \tag{11}$$

sont réalisées identiquement en s, $a < s < b$;

2° *pour* $s = 0$ *on a*

$$x_1(0) = x_1^{(0)}, \ \ldots, \ x_m(0) = x_m^{(0)}. \ \square$$

On appliquera ce théorème aux relations (7) qui, pour les fonctions données $k_1, \ldots, k_{n-1}$, sont des équations de la forme (11) pour $m = n^2$ coordonnées des vecteurs de base $\mathbf{t}_1, \ldots, \mathbf{t}_n$. En vertu du théorème d'EUS, il existe sur $]a, b[$ une famille de fonctions vectorielles $\mathbf{t}_1(s), \ldots, \mathbf{t}_{n-1}(s)$, $a < s < b$, et une seule, telle que

1° pour tout s on a les relations (7);

2° pour $s = 0$

$$\mathbf{t}_1(0) = \mathbf{i}_1, \ \ldots, \ \mathbf{t}_n(0) = \mathbf{i}_n. \tag{12}$$

E t a p e 2. Considérons les produits scalaires $\mathbf{t}_i\mathbf{t}_j$, $i, j = 1, \ldots, n$. En vertu de (7) on a pour ces produits

$$(\mathbf{t}_i\mathbf{t}_j)^{\cdot} = \dot{\mathbf{t}}_i\mathbf{t}_j + \mathbf{t}_i\dot{\mathbf{t}}_j = (-k_{i-1}\mathbf{t}_{i-1} + k_i\mathbf{t}_{i+1})\,\mathbf{t}_j + \mathbf{t}_i\,(-k_{j-1}\mathbf{t}_{j-1} + k_j\mathbf{t}_{j+1})$$

(on convient que $\mathbf{t}_0 = 0$ et $\mathbf{t}_{n+1} = 0$), i.e. on a obtenu des égalités

$$(13) \qquad (\mathbf{t}_i\mathbf{t}_j)^{\cdot} = -k_{i-1}(\mathbf{t}_{i-1}\mathbf{t}_j) + k_i(\mathbf{t}_{i+1}\mathbf{t}_j) - k_{j-1}(\mathbf{t}_i\mathbf{t}_{j-1}) + k_j(\mathbf{t}_i\mathbf{t}_{j+1})$$

qui peuvent être traitées comme des équations (11) pour $m = \frac{n(n+1)}{2}$ fonctions $\mathbf{t}_i\mathbf{t}_j$. Donc, en vertu du théorème d'EUS, il existe une seule famille de fonctions $\mathbf{t}_i\mathbf{t}_j$, $i, j = 1, \ldots, n$, égales à $\delta_{ij} = \mathbf{i}_i\mathbf{i}_j$ pour $s = 0$ (i.e. à 0 si $i \neq j$ et à 1 si $i = j$).

D'autre part, une vérification immédiate montre que les équations (13) sont satisfaites par des fonctions $\mathbf{t}_i\mathbf{t}_j$ identiquement égales à δ_{ij}. (En effet, pour $i \neq j - 1, j + 1$, tous les termes de la somme $-k_{i-1}\delta_{i-1,j} + k_i\delta_{i+1,j} - k_{j-1}\delta_{i,j-1} + k_j\delta_{i,j+1}$ sont nuls et pour $i = j - 1$, $j + 1$, cette somme ne contient que deux termes non nuls mais qui se détruisent.) Donc, d'après le théorème d'EUS, pour tout s on a les égalités $\mathbf{t}_i\mathbf{t}_j = \delta_{ij}$, $i, j = 1, \ldots, n$, qui expriment que les vecteurs $\mathbf{t}_1, \ldots, \mathbf{t}_n$ forment une base orthonormale pour tout s, $a < s < b$.

Cette base étant confondue pour $s = 0$ avec la base directe $\mathbf{i}_1, \ldots, \mathbf{i}_n$, on en déduit que la base $\mathbf{t}_1, \ldots, \mathbf{t}_n$ est directe pour tout s, $a < s < b$.

E t a p e 3. Prenons les dérivées successives du vecteur $\mathbf{t}_1$:

$$(14) \qquad \mathbf{t}_1, \dot{\mathbf{t}}_1, \ddot{\mathbf{t}}_1, \ldots, \overset{(n-1)}{\mathbf{t}_1},$$

et appliquons-leur le processus d'orthogonalisation de Gram-Schmidt. On ne touchera pas au vecteur $\mathbf{t}_1$, puisque c'est déjà un vecteur unitaire. Le vecteur $\dot{\mathbf{t}}_1$ est orthogonal à $\mathbf{t}_1$ d'après le lemme 2 de la leçon 1, donc on doit seulement le normer. D'autre part, puisque d'après ce qui a été prouvé le vecteur $\mathbf{t}_2$ est un vecteur unitaire et par hypothèse $k_1 > 0$, la première des relations (7) nous dit que $|\mathbf{t}_1| = k_1$. Donc, au deuxième pas du processus d'orthogonalisation on obtient le vecteur

$$\mathbf{t}_2 = \frac{\dot{\mathbf{t}}_1}{k_1}.$$

Au troisième pas il faut envisager le vecteur

$$\ddot{\mathbf{t}}_1 = (k_1\mathbf{t}_2)^{\cdot} = \dot{k}_1\mathbf{t}_2 + k_1\dot{\mathbf{t}}_2 = -k_1^2\mathbf{t}_1 + \dot{k}_1\mathbf{t}_2 + k_1k_2\mathbf{t}_3.$$

lui soustraire une combinaison linéaire des vecteurs $\mathbf{t}_1$ et $\mathbf{t}_2$ de façon à obtenir un vecteur qui leur soit orthogonal à tous et enfin le normer. Mais comme les vecteurs $\mathbf{t}_1$, $\mathbf{t}_2$, $\mathbf{t}_3$ forment, en vertu de ce qui a été

prouvé, un système orthonormal et que par hypothèse $k_1 k_2 > 0$, cette procédure nous donne de toute évidence le vecteur $\mathbf{t}_3$.

Notre raisonnement étant de toute évidence général, à chaque pas de la procédure d'orthogonalisation on obtiendra un vecteur $\mathbf{t}_i$, $i = 1, \ldots, n-1$, convenable. Ceci prouve que le système de vecteurs $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_{n-1}$ est un système orthonormal obtenu d'une façon unique à partir du système (14) par le procédé d'orthogonalisation de Gram-Schmidt.

E t a p e 4. Soit

$$\mathbf{r}(s) = \int_0^s \mathbf{t}_1(s)\,ds, \quad a < s < b. \tag{15}$$

Alors $\mathbf{r}(0) = 0$ et $\dot{\mathbf{r}}(s) = \mathbf{t}_1(s)$, i.e. la courbe $\mathbf{r} = \mathbf{r}(s)$, $a < s < b$, passe par le point O pour $s = 0$ et le vecteur $\mathbf{t}_1(s)$ lui est tangent pour tout s. Or, pour chaque courbe, les $n-1$ premiers vecteurs du repère mobile sont des vecteurs obtenus à partir des $n-1$ premières dérivées du vecteur tangent par le procédé d'orthogonalisation de Gram-Schmidt. Donc, d'après ce qui a été prouvé plus haut, ces vecteurs sont confondus avec les vecteurs $\mathbf{t}_1, \ldots, \mathbf{t}_{n-1}$.

S'agissant du dernier vecteur du repère mobile, c'est le seul vecteur unitaire qui forme une base directe avec les $n-1$ premiers. La base $\mathbf{t}_1, \ldots, \mathbf{t}_n$ étant directe, ce vecteur est nécessairement le vecteur $\mathbf{t}_n$.

Ainsi on a prouvé que pour tout s les vecteurs $\mathbf{t}_1(s), \ldots, \mathbf{t}_n(s)$ forment une base mobile de la courbe $\mathbf{r} = \mathbf{r}(s)$. Puisque ces vecteurs sont justiciables des formules de Frenet (7), les fonctions $k_i(s)$, $i = 1, \ldots, n-1$, qui y figurent, doivent être les courbures de la courbe $\mathbf{r} = \mathbf{r}(s)$.

Ceci prouve entièrement l'existence d'une courbe $\mathbf{r} = \mathbf{r}(s)$ jouissant des propriétés 1° et 2°.

L'unicité de cette courbe découle du fait qu'aux termes du théorème d'EUS la base mobile $\mathbf{t}_1(s), \ldots, \mathbf{t}_n(s)$ est définie de façon unique par les équations (7) et les conditions initiales (12), et le rayon vecteur $\mathbf{r}(s)$, de façon unique par la relation $\dot{\mathbf{r}}(s) = \mathbf{t}_1(s)$ (formule (15)) et la condition initiale $\mathbf{r}(0) = 0$. □

LEÇON 3

Surfaces élémentaires et leurs paramétrisations. — Exemples de surfaces. — Plan tangent et sous-espace tangent. — Applications différentiables de surfaces et leurs différentielles. — Difféomorphismes de surfaces. — Première forme quadratique d'une surface. — Isométries. — Premier paramètre différentiel de Beltrami. — Exemples de calcul des premières formes quadratiques. — Surfaces développables.

La définition rigoureuse de la notion de surface se fait par analogie à celle d'une ligne et se heurte à de plus grandes difficultés. C'est pourquoi on se bornera pour l'instant à l'analogue de la notion d'arc régulier simple ouvert (ce qui ne nous empêchera pas d'envisager des surfaces plus générales dans des exemples concrets).

Pour introduire cet analogue on commencera par une application continue arbitraire de la forme

$$\gamma : U \to \mathcal{A}, \tag{1}$$

où $\mathcal{A}$ est comme toujours un espace euclidien (ou plus simplement affine) de dimension $n \geqslant 3$, et U un sous-ensemble ouvert et *convexe* (i.e. contenant tout couple de points avec le segment de droite qui les joint) du plan arithmétique $\mathbb{R}^2$ (l'analogue bidimensionnel d'un intervalle $I =]a, b[$). Lorsque $\mathcal{A}$ est rapporté à une origine O, l'application (1) est définie par une fonction vectorielle continue

$$\mathbf{r} = \mathbf{r}(u, v), \ (u, v) \in U, \tag{2}$$

à valeurs dans un espace vectoriel associé $\mathcal{V}$; lorsque $\mathcal{V}$ est muni d'une base $\mathbf{e}_1, \ldots, \mathbf{e}_n$, l'application (1) est définie par n fonctions numériques continues

$$x^1 = x^1(u, v), \ \ldots, \ x^n = x^n(u, v), \tag{3}$$

les composantes du vecteur $\mathbf{r}(u, v)$ dans la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

On dit qu'une application (1) est *de classe* C^r, où r est un entier naturel ou le symbole ∞, si chaque fonction (3) — ou, ce qui est équivalent, la fonction vectorielle (2) — possède des dérivées partielles continues de tous les ordres $\leqslant r$ (on rappelle que pour $r = \infty$ cela signifie qu'existent les dérivées partielles de tous les ordres). Dans la suite on admettra que le nombre r est fixé une fois pour toutes et qu'il est suffisamment grand. [Dans cette leçon, tout nombre $r \geqslant 1$ nous conviendra, mais à la leçon 5 il faudra que $r \geqslant 3$.]

En particulier, pour l'application dérivable (1) sont définies les dérivées partielles

$$\mathbf{r}_u = \frac{\partial \mathbf{r}(u, v)}{\partial u}, \quad \mathbf{r}_v = \frac{\partial \mathbf{r}(u, v)}{\partial v} \tag{4}$$

de la fonction vectorielle (2). Une application dérivable (1) est dite *régulière* si les dérivées partielles (4) sont linéairement indépendantes en chaque point $(u, v) \in U$.

Définition 1. On dit qu'une application (1) est une *paramétrisation* si elle est

1° dérivable,

2° régulière,

3° monéomorphe (i.e. est injective et possède la propriété suivante: si une suite de points $\gamma(u_n, v_n)$, $(u_n, v_n) \in U$, de l'espace $\mathcal{A}$ converge vers un point de la forme $\gamma(a, b)$, où $(a, b) \in U$, la suite de points $(u_n, v_n) \in U$ converge aussi (nécessairement vers le point (a, b) en raison de la continuité)).

Exercice 1. Montrer que, si une application (1) est dérivable et régulière, tout point $(u_0, v_0) \in U$ possède un voisinage $V \subset U$ sur lequel cette application est monéomorphe (est une paramétrisation).

Définition 2. On dit qu'un sous-ensemble $\mathcal{X}$ de $\mathcal{A}$ est une *surface élémentaire* s'il existe une paramétrisation $\gamma: U \to \mathcal{A}$ (appelée *paramétrisation de la surface* $\mathcal{X}$) telle que $\gamma(U) = \mathcal{X}$. On dit aussi que $\mathcal{X}$ est le *support* de la paramétrisation γ.

Les surfaces élémentaires sont les analogues bidimensionnels des arcs réguliers simples (leurs paramétrisations, les analogues des courbes régulières simples).

Remarque 1. Dans une autre terminologie dont on se servira parfois sans le mentionner expressément, on appellera surfaces élémentaires les paramétrisations (1) elles-mêmes. A la leçon 1 nous avons évité cette homonimie en faisant une distinction entre les courbes et les lignes. Dans le cas bidimensionnel il n'existe malheureusement pas de couple de termes analogue.

Puisqu'on n'étudiera pas de surfaces autres que les surfaces élémentaires, on omettra dans la suite le qualificatif « élémentaire ».

La paramétrisation $\gamma: U \to \mathcal{A}$ d'une surface $\mathcal{X}$ étant une application injective, pour tout point $p \in \mathcal{X}$ il existe des nombres uniques u et v tels que $(u, v) \in U$ et $\gamma(u, v) = p$. Ces nombres s'appellent *coordonnées* du point p pour cette paramétrisation. Ces coordonnées sont traditionnellement affectées des épithètes *curvilignes* ou *locales*, épithètes qui en principe sont superflues, car ces coordonnées sont les seules qui soient généralement envisagées sur les surfaces.

Par abus de langage on dira aussi parfois que les nombres u et v sont des coordonnées *sur la surface* (1), en souscrivant à la tendance générale à confondre les surfaces et leurs paramétrisations.

Chaque courbe de U d'équations paramétriques

$$u = u(t), \quad v = v(t), \quad t \in I, \tag{5}$$

est envoyée par la paramétrisation (1) de la surface $\mathcal{X}$ dans une courbe

$$\mathbf{r} = \mathbf{r}(u(t), v(t)), \quad t \in I, \tag{6}$$

de l'espace $\mathcal{A}$. On dit de la courbe (6) qu'elle est *située* sur la surface $\mathcal{A}$ et que les équations (5) sont ses *équations paramétriques par rapport aux coordonnées u et v.*

En particulier, les courbes $u =$ const et $v =$ const (qui sont les images des lignes de coordonnées dans U) s'appellent *lignes de coordonnées sur la surface $\mathcal{X}$*, et leur ensemble, *réseau de coordonnées.*

Pour tous sous-ensembles ouverts U, $U^* \subset \mathbb{R}^2$, toute application

$$\varphi: U^* \to U \tag{7}$$

est définie par un couple de fonctions

$$u = u(u^*, v^*), \quad v = v(u^*, v^*), \quad (u^*, v^*) \in U^*, \tag{8}$$

telles que pour tout point $(u^*, v^*) \in U^*$ le point $(u, v) = (u(u^*, v^*), v(u^*, v^*))$ appartient à U. On dit qu'une application (7) est *de classe C^r*, où r est un entier naturel ou le symbole ∞, si les fonctions (8) possèdent, pour $r \neq \infty$, des dérivées partielles continues de tout ordre $\leqslant r$. Une application différentiable (7) est un *difféomorphisme* si elle est bijective et si sa réciproque

$$\varphi^{-1}: U \to U^*$$

est différentiable aussi.

Deux paramétrisations

$$\gamma: U \to \mathcal{A} \quad \text{et} \quad \gamma^*: U^* \to \mathcal{A} \tag{9}$$

sont *équivalentes* s'il existe un difféomorphisme (7) tel que

$$\gamma^* = \gamma \circ \varphi. \tag{10}$$

Puisque la donnée d'une paramétrisation de la surface $\mathcal{X}$ équivaut à la donnée de coordonnées curvilignes sur $\mathcal{X}$, on dit aussi du difféomorphisme (7) qu'il réalise un *changement de coordonnées* sur $\mathcal{X}$ (ou définit un *passage* des coordonnées u, v aux coordonnées u^*, v^*).

Exercice 2. Montrer que la relation (10) est une relation d'équivalence (i.e. est réflexive, symétrique et transitive) et qu'on peut donc parler des classes d'équivalence de paramétrisations.

Il est clair que *les paramétrisations équivalentes possèdent le même support.* Réciproquement, on peut montrer sans peine que *des paramétrisations* (9) *possédant le même support sont équivalentes.* Ceci

exprime que les surfaces se trouvent en correspondance bijective avec les classes d'équivalence de leurs paramétrisations et peuvent donc leur être identifiées.

Exercice 3. Prouver la dernière assertion. (On remarquera que les conditions 1°, 2° et 3° de la définition 1 sont essentielles à sa validité. Comparer avec la démonstration de la proposition 1 de la leçon 1.)

A la leçon 15 on établira une proposition générale dont cette assertion (ainsi que la proposition 1 de la leçon 1) est un cas particulier.

* * *

Exemples de surfaces.

Pour la suggestion, on se bornera aux surfaces de l'espace euclidien rapporté aux coordonnées rectangulaires x, y, z.

Exemple 1. Les équations

$$x = R \cos u, \quad y = R \sin u, \quad z = v \tag{11}$$

définissent un cylindre circulaire droit. Ce cylindre n'est pas une surface élémentaire au sens de la définition 2, car pour $-\infty < u < +\infty$ tout point de ce cylindre est recouvert une infinité (dénombrable) de fois par les points du plan $\mathbb{R}^2$. Pour obtenir une surface élémentaire il faut couper le cylindre le long d'une génératrice, i.e. assujettir le paramètre u à la double inégalité $0 < u < 2\pi$ dans les équations (11). Le cylindre sera alors recouvert par deux cylindres ainsi coupés.

Le réseau de coordonnées du cylindre (11) est composé de droites verticales $u = \text{const}$ et de cercles horizontaux $v = \text{const}$.

Exemple 2. Soit $x = x(v)$, $z = z(v)$ une courbe régulière simple du plan Oxz ne coupant pas l'axe Oz. La surface de paramétrisation

$$x = x(v) \cos u, \quad y = x(v) \sin u, \quad z = z(v) \tag{12}$$

s'appelle *surface de révolution*, et la courbe $x = x(v)$, $z = z(v)$ son *profil*. La surface (12) s'obtient par rotation de son profil autour de l'axe Oz.

La régularité de la paramétrisation (12), i.e. l'indépendance linéaire des vecteurs

$$\mathbf{r}_u = (-x(v) \sin u,\ x(v) \cos u,\ 0),$$
$$\mathbf{r}_v = (x'(v) \cos u,\ x'(v) \sin u,\ z'(v)),$$

résulte de celle du profil (i.e. de la condition $x'(v)^2 + z'(v)^2 = 1$) et du fait que le profil ne coupe pas l'axe de rotation Oz (i.e. du fait que $x(v) \neq 0$).

Le réseau de coordonnées de la surface (12) est constitué de courbes obtenues par rotation du profil autour de l'axe Oz (les *méridiens*)

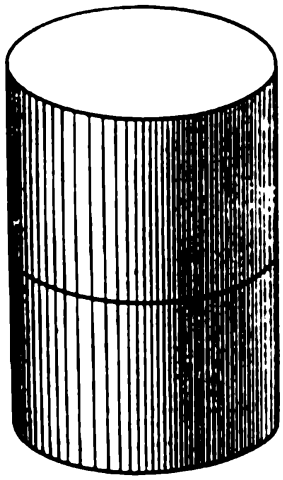
et de cercles perpendiculaires à ces courbes (les *parallèles*). Pour rendre la surface (12) élémentaire il faut la couper le long d'un méridien.

Le cylindre est une surface de révolution dont le profil est une droite $x = R$, $z = v$.

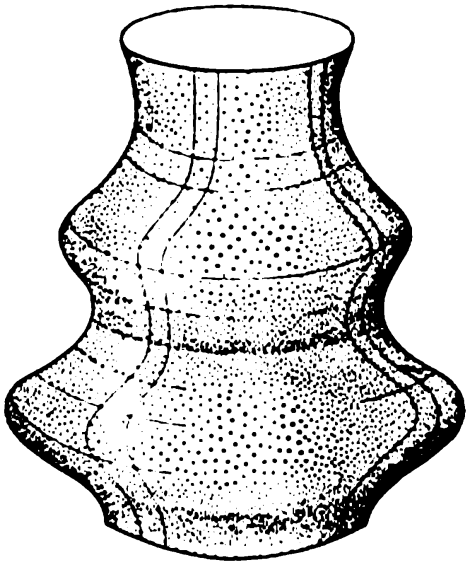
La surface de révolution de profil $x = R \cos v$, $z = R \sin v$ (un cercle) est la sphère

$$x = R \cos v \cos u, \quad y = R \cos v \sin u, \quad z = R \sin v$$

de rayon R et de centre O. Les coordonnées u et v de la sphère sont les notoires « coordonnées géographiques »: la longitude et la latitude, et les courbes de coordonnées, les méridiens et les parallèles.



Cylindre circulaire



Surface de révolution

En toute rigueur, le profil de la sphère est le demi-cercle $-\pi < v < +\pi$ (ce qui exclut les pôles). D'autre part, pour obtenir une surface élémentaire il faut éliminer un méridien (« la ligne de changement des dates »).

Exemple 3. Une surface $\mathbf{r} = \mathbf{r}(u, v)$ s'appelle *surface réglée* si

$$\mathbf{r}(u, v) = \boldsymbol{\rho}(u) + v\mathbf{a}(u), \tag{13}$$

où $\boldsymbol{\rho}(u)$ et $\mathbf{a}(u)$ sont des fonctions vectorielles telles que les vecteurs $\boldsymbol{\rho}'(u) + v\mathbf{a}'(u)$ et $\mathbf{a}(u)$ sont linéairement indépendants pour les u et v envisagés (de sorte que, en particulier, $\mathbf{a}(u) \neq 0$ pour tout u). A noter que cette indépendance linéaire assure la régularité. La courbe de coordonnées $u = u_0 = \text{const}$ est la droite de vecteur directeur $\mathbf{a}(u_0)$ passant par le point de rayon vecteur $\boldsymbol{\rho}(u_0)$. Donc, de façon imagée, une surface réglée est balayée par une droite mobile dans l'espace. Comparer avec la définition 1 de la leçon I.23.

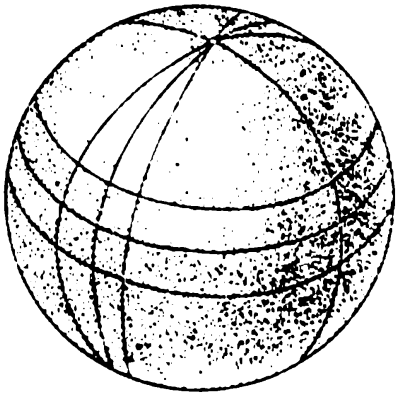
Il est clair qu'on peut, sans perdre en généralité, admettre que le vecteur $\mathbf{a}(u)$ est unitaire :

$$|\mathbf{a}(u)| = 1 \quad \text{pour tout } u.$$

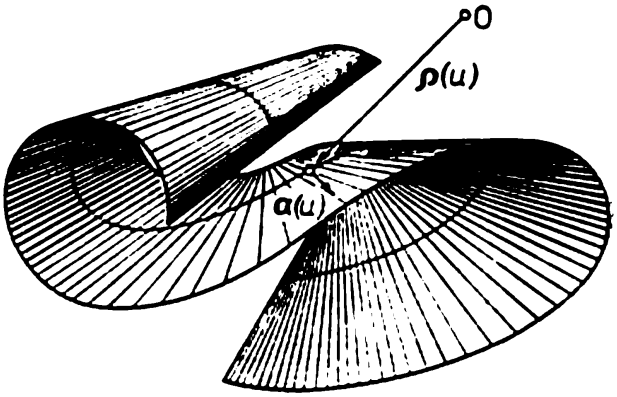
Si $\rho'(u) = 0$ pout tout u, i.e. $\rho(u) = \text{const}$, par transport de l'origine des coordonnées on obtient, au lieu de (13), l'équation

$$\mathbf{r} = v\mathbf{a}(u) \tag{14}$$

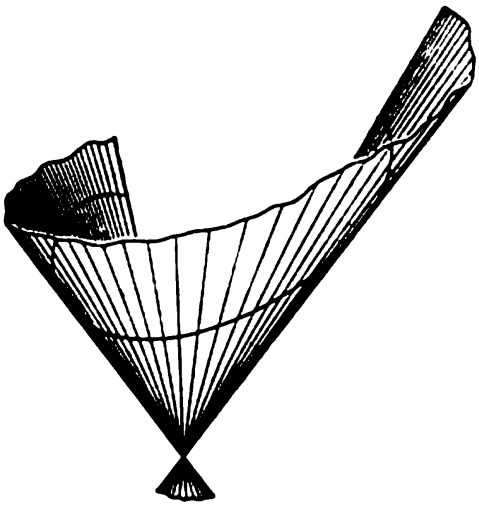
qui définit un cône dont la directrice est une courbe gauche régulière $\mathbf{r} = \mathbf{a}(u)$. [Dans l'équation (14) il faut de toute évidence con-



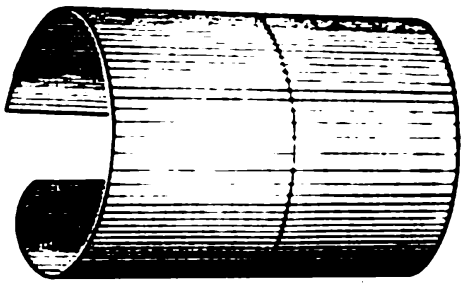
Sphère



Surface réglée



Cône



Cylindre

sidérer que $v > 0$ ou $v < 0$ (car le point $v = 0$ est un point singulier et partage le cône en deux nappes). Lorsque les génératrices du cône coupent une directrice en plusieurs points, il faut astreindre v à des contraintes supplémentaires.]

Si $\mathbf{a}'(u) = 0$ pour tout u, i.e. $\mathbf{a}(u) = \text{const}$, la surface (14) est un cylindre dont la directrice est une courbe $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$ (généralement gauche).

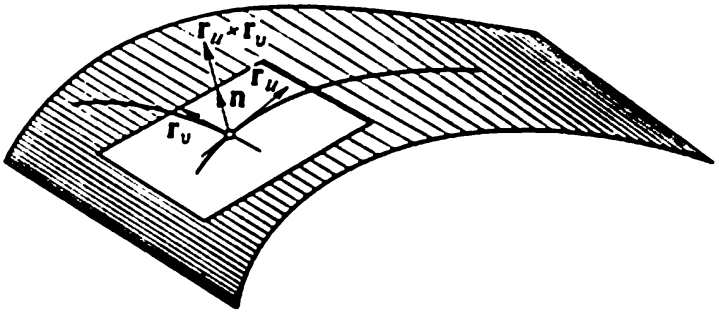
Si le vecteur $\boldsymbol{\rho}'$ n'est pas identiquement nul, on peut admettre, quitte à passer à un plus petit domaine dans $\mathbb{R}^2$, que $\boldsymbol{\rho}'(u) \neq 0$ pour tout u. Alors $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$ sera une courbe régulière gauche et on peut admettre que u est un paramètre naturel (une longueur d'arc). Le cône (14) peut aussi être défini par une équation (13) avec $\boldsymbol{\rho}'(u) \neq 0$. Pour cela il suffit de poser $\boldsymbol{\rho}(u) = \mathbf{a}(u)$ dans (13) (sous réserve bien sûr que $\mathbf{a}'(u) \neq 0$).

Si $\mathbf{a}(u)$ est un vecteur tangent $\boldsymbol{\tau}(u)$ à la courbe $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$, la surface (13) s'appelle *surface des tangentes*. On définit de façon analogue la *surface des normales principales* et la *surface des binormales*.

Signalons que tous les points de la courbe $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$ sont des points singuliers de la surface des tangentes. (Ces points en lesquels la condition de régularité est violée constituent l'*arête de rebroussement* de la surface des tangentes.)

* * *

Supposons de nouveau que $\mathscr{X}$ est une surface d'un espace affine $\mathscr{A}$ à n dimensions de paramétrisation $\mathbf{r} = \mathbf{r}(u, v)$, et soient p_0



Plan tangent

un point arbitraire de $\mathscr{X}$ et $\mathbf{r}_0 = \mathbf{r}(u_0, v_0)$, le rayon vecteur de p_0. En vertu de la condition de régularité, les valeurs

$$\mathbf{r}_{u_0} = \frac{\partial \mathbf{r}}{\partial u}(u_0, v_0), \qquad \mathbf{r}_{v_0} = \frac{\partial \mathbf{r}}{\partial v}(u_0, v_0)$$

des dérivées partielles de la fonction vectorielle $\mathbf{r} = \mathbf{r}(u, v)$ au point (u_0, v_0) sont linéairement indépendantes, i.e. le bivecteur $\mathbf{r}_{u_0} \wedge \mathbf{r}_{v_0}$ est non nul. Donc, dans l'espace $\mathscr{A}$ est défini un plan de bivecteur directeur $\mathbf{r}_{u_0} \wedge \mathbf{r}_{v_0}$ passant par p_0. L'équation paramétri-

que vectorielle est de la forme

$$\mathbf{r} = \mathbf{r}_0 + a\mathbf{r}_{u_0} + b\mathbf{r}_{v_0}, \tag{15}$$

où a et b sont des paramètres.

Définition 3. Le plan (15) s'appelle *plan tangent* à la surface $\mathscr{X}$ en p_0. Le sous-espace correspondant de l'espace vectoriel associé $\mathscr{V}$ (qui est composé des vecteurs de la forme $a\mathbf{r}_{u_0} + b\mathbf{r}_{v_0}$ et désigné par $\mathrm{T}_{p_0}\mathscr{X}$) s'appelle *sous-espace tangent* et ses vecteurs, *vecteurs tangents* à $\mathscr{X}$ en p_0. Pour souligner que le sous-espace tangent est bidimensionnel, on l'appelle souvent *plan tangent*.

Cette terminologie est justifiée par le fait que le vecteur tangent

$$\mathbf{r}'(t_0) = u'(t_0)\,\mathbf{r}_{u_0} + v'(t_0)\,\mathbf{r}_{v_0} \tag{16}$$

à une courbe (5) de $\mathscr{X}$, passant (disons pour $t = t_0$) par un point p_0, appartient au sous-espace tangent $\mathrm{T}_{p_0}\mathscr{X}$ et, réciproquement, tout vecteur $a\mathbf{r}_u + b\mathbf{r}_v$ de $\mathrm{T}_{p_0}\mathscr{X}$ peut être mis sous la forme (16) (il suffit d'envisager une courbe d'équations paramétriques $u = u_0 + at$, $v = v_0 + bt$, $t \in I$, où I est un intervalle de l'axe t tel que $(u_0 + at, v_0 + bt) \in U$ pour tout $t \in I$). Donc, *les vecteurs tangents à une surface sont des vecteurs tangents à des courbes de cette surface.*

La relation (10) qui définit un changement de coordonnées sur la surface $\mathscr{X}$ s'écrit en termes de rayons vecteurs:

$$\mathbf{r}^*(u^*, v^*) = \mathbf{r}(u(u^*, v^*), v(u^*, v^*)).$$

La dérivation de cette formule nous donne

$$\begin{aligned} \mathbf{r}^*_{u^*} &= \frac{\partial u}{\partial u^*}\,\mathbf{r}_u + \frac{\partial v}{\partial u^*}\,\mathbf{r}_v, \\ \mathbf{r}^*_{v^*} &= \frac{\partial u}{\partial v^*}\,\mathbf{r}_u + \frac{\partial v}{\partial v^*}\,\mathbf{r}_v. \end{aligned} \tag{17}$$

Des relations (17) il s'ensuit que *les vecteurs* $\mathbf{r}^*_{u^*}$ *et* $\mathbf{r}^*_{v^*}$ *sont linéairement équivalents aux vecteurs* $\mathbf{r}_u$ *et* $\mathbf{r}_v$, donc engendrent le même sous-espace. Ceci prouve que *pour tout point* $p \in \mathscr{X}$ *l'espace vectoriel* $\mathrm{T}_p\mathscr{X}$ *est défini de façon unique* (i.e. ne dépend pas du choix de la paramétrisation $\mathbf{r} = \mathbf{r}(u, v)$). Si l'on change de paramétrisation dans l'espace $\mathrm{T}_p\mathscr{X}$, seule change (à l'aide des formules (17)) la base $\mathbf{r}_u$, $\mathbf{r}_v$.

Les vecteurs de l'espace $\mathrm{T}_p\mathscr{X}$ sont souvent désignés par $\mathrm{d}\mathbf{r}$ et leurs coordonnées (dans la base $\mathbf{r}_u$, $\mathbf{r}_v$) par $\mathrm{d}u$ et $\mathrm{d}v$ (ces coordonnées sont écrites à droite des vecteurs $\mathbf{r}_u$ et $\mathbf{r}_v$). Dans ces notations on a donc

$$\mathrm{d}\mathbf{r} = \mathbf{r}_u\,\mathrm{d}u + \mathbf{r}_v\,\mathrm{d}v \tag{18}$$

pour tout vecteur $\mathrm{d}\mathbf{r}$ de l'espace $\mathrm{T}_p\mathscr{X}$. Ceci étant, il s'ensuit immédiatement des formules (17) que les coordonnées $\mathrm{d}u$ et $\mathrm{d}v$ sont liées

aux coordonnées du^*, dv^* dans la base $\mathbf{r}^*_{u^*}$, $\mathbf{r}^*_{v^*}$ par les relations

$$(19)\qquad \begin{aligned} du &= \frac{\partial u}{\partial u^*}\,du^* + \frac{\partial u}{\partial v^*}\,dv^*, \\ dv &= \frac{\partial v}{\partial u^*}\,du^* + \frac{\partial v}{\partial v^*}\,dv^*, \end{aligned}$$

qui formellement sont confondues avec les classiques formules des différentielles en analyse (ce qui plaide en faveur des notations (18)).

* * *

Considérons maintenant, parallèlement à une surface $\mathscr{X}$ de paramétrisation $\gamma: U \to \mathscr{A}$ définie par une fonction vectorielle $\mathbf{r} = \mathbf{r}(u, v)$, $(u, v) \in U$, une autre surface élémentaire $\hat{\mathscr{X}}$ de paramétrisation $\hat{\gamma}: \hat{U} \to \mathscr{A}$ définie par une fonction vectorielle $\mathbf{r} = \hat{\mathbf{r}}(\hat{u}, \hat{v})$, $(\hat{u}, \hat{v}) \in \hat{U}$.

Les applications γ et $\hat{\gamma}$ étant injectives, chaque application $f: \mathscr{X} \to \hat{\mathscr{X}}$ définit de façon unique une application $\bar{f}: U \to \hat{U}$ satisfaisant à la relation

$$(20)\qquad f \circ \gamma = \hat{\gamma} \circ \bar{f}$$

et définissant de façon unique l'application f. De façon imagée, la relation (20) exprime que dans le diagramme

$$(21)\qquad \begin{tikzcd} \mathscr{X} \arrow[r, "f"] & \hat{\mathscr{X}} \\ U \arrow[u, "\gamma"] \arrow[r, "\bar{f}"] & \hat{U} \arrow[u, "\hat{\gamma}"'] \end{tikzcd}$$

on aboutit au même résultat quel que soit celui des deux chemins pris pour aller de U en $\hat{\mathscr{X}}$. [Les diagrammes jouissant de cette propriété sont dits *commutatifs*.]

On dit de l'application $\bar{f}$ qu'elle *représente l'application f dans les paramétrisations γ et $\hat{\gamma}$*, et des fonctions

$$(22)\qquad \hat{u} = \hat{u}(u, v), \quad \hat{v} = \hat{v}(u, v), \quad (u, v) \in U,$$

définissant l'application $\bar{f}$ qu'elles *définissent l'application f dans les coordonnées u, v et u^*, v^**.

L'application f est *différentiable* si l'application $\bar{f}$ l'est, i.e. si les fonctions (22) le sont. Il est clair que cette définition est intrinsè-

que (si une application $\bar{f}$ est différentiable pour un choix des paramétrisations γ et $\hat{\gamma}$, elle le sera pour tout autre choix).

A toute application différentiable $f: \mathscr{X} \to \hat{\mathscr{X}}$ et tout point $p \in \mathscr{X}$ associons une application linéaire

$$\mathrm{T}_p\mathscr{X} \to \mathrm{T}_{\hat{p}}\hat{\mathscr{X}}, \quad \hat{p} = f(p), \tag{23}$$

qui envoie un vecteur (18) de $\mathrm{T}_p\mathscr{X}$ dans un vecteur

$$\mathrm{d}\hat{\mathbf{r}} = \hat{\mathbf{r}}_{\hat{u}}\,\mathrm{d}\hat{u} + \hat{\mathbf{r}}_{\hat{v}}\,\mathrm{d}\hat{v}$$

de $\mathrm{T}_{\hat{p}}\hat{\mathscr{X}}$, où conformément aux formules de calcul différentiel

$$\begin{aligned} \mathrm{d}\hat{u} &= \frac{\partial \hat{u}}{\partial u}\,\mathrm{d}u + \frac{\partial \hat{u}}{\partial v}\,\mathrm{d}v, \\ \mathrm{d}\hat{v} &= \frac{\partial \hat{v}}{\partial u}\,\mathrm{d}u + \frac{\partial \hat{v}}{\partial v}\,\mathrm{d}v. \end{aligned} \tag{24}$$

L'application (23) est définie pour les paramétrisations γ, $\hat{\gamma}$ des surfaces $\mathscr{X}$, $\hat{\mathscr{X}}$, d'où la question légitime de son indépendance par rapport au choix de ces paramétrisations.

Supposons par exemple que nous avons remplacé la paramétrisation γ par une autre paramétrisation $\gamma^*: U^* \to \mathscr{A}$ de la surface $\mathscr{X}$. Par définition, $\gamma^* = \gamma \circ \varphi$, où $\varphi: U^* \to U$ est un difféomorphisme défini par des fonctions

$$u = u(u^*, v^*), \quad v = v(u^*, v^*).$$

Ceci étant, les bases respectives $\mathbf{r}_u$, $\mathbf{r}_v$ et $\mathbf{r}_{u^*}$, $\mathbf{r}_{v^*}$ de l'espace $\mathrm{T}_p\mathscr{X}$ seront justiciables des formules (17), et les coordonnées respectives $\mathrm{d}u$, $\mathrm{d}v$ et $\mathrm{d}u^*$, $\mathrm{d}v^*$ des vecteurs tangents, des formules (19). En coordonnées u^*, v^* et $\hat{u}$, $\hat{v}$, l'application f est définie de toute évidence par les fonctions

$$\hat{u}^*(u^*, v^*) = \hat{u}(u(u^*, v^*), v(u^*, v^*)),$$

$$\hat{v}^*(u^*, v^*) = \hat{v}(u(u^*, v^*), v(u^*, v^*)),$$

donc l'image du vecteur $\mathrm{d}\mathbf{r} \in \mathrm{T}_p\mathscr{X}$ de coordonnées $\mathrm{d}u^*$, $\mathrm{d}v^*$ dans la base $\mathbf{r}_{u^*}$, $\mathbf{r}_{v^*}$ par l'application (23) construite à l'aide des paramétrisations γ^*, $\hat{\gamma}$ sera le vecteur $\mathrm{d}\hat{\mathbf{r}}^* \in \mathrm{T}_{\hat{p}}\hat{\mathscr{X}}$ de coordonnées

$$\mathrm{d}\hat{u}^* = \frac{\partial \hat{u}^*}{\partial u^*}\,\mathrm{d}u^* + \frac{\partial \hat{u}^*}{\partial v^*}\,\mathrm{d}v^*,$$

$$\mathrm{d}\hat{v}^* = \frac{\partial \hat{v}^*}{\partial u^*}\,\mathrm{d}u^* + \frac{\partial \hat{v}^*}{\partial v^*}\,\mathrm{d}v^*$$

dans la base $\hat{\mathbf{r}}_{\hat{u}}$, $\hat{\mathbf{r}}_{\hat{v}}$. D'autre part, en portant les expressions (19) des coordonnées du et dv dans les formules (24) et en tenant compte du fait qu'en vertu des règles de dérivation des fonctions composées

$$\frac{\partial \hat{u}}{\partial u}\frac{\partial u}{\partial u^*}+\frac{\partial \hat{u}}{\partial v}\frac{\partial v}{\partial u^*}=\frac{\partial \hat{u}^*}{\partial u^*}, \quad \frac{\partial \hat{u}}{\partial u}\frac{\partial u}{\partial v^*}+\frac{\partial \hat{u}}{\partial v}\frac{\partial v}{\partial v^*}=\frac{\partial \hat{u}^*}{\partial v^*},$$

$$\frac{\partial \hat{v}}{\partial u}\frac{\partial u}{\partial u^*}+\frac{\partial \hat{v}}{\partial v}\frac{\partial v}{\partial u^*}=\frac{\partial \hat{v}^*}{\partial u^*}, \quad \frac{\partial \hat{v}}{\partial u}\frac{\partial u}{\partial v^*}+\frac{\partial \hat{v}}{\partial v}\frac{\partial v}{\partial v^*}=\frac{\partial \hat{v}^*}{\partial v^*},$$

on trouve aussitôt que $\mathrm{d}\hat{u} = \mathrm{d}\hat{u}^*$ et $\mathrm{d}\hat{v} = \mathrm{d}\hat{v}^*$, i.e. que $\mathrm{d}\hat{\mathbf{r}} = \mathrm{d}\hat{\mathbf{r}}^*$. Ceci montre que le vecteur $\mathrm{d}\hat{\mathbf{r}}$ ne dépend pas du choix de la paramétrisation γ. On démontre de façon analogue (faites-le à titre d'exercice) que ce vecteur ne dépend pas non plus du choix de la paramétrisation $\hat{\gamma}$. Donc, *l'application* (23) *est définie de façon intrinsèque.*

Définition 4. L'application (23) s'appelle *différentielle* de l'application f en p (ou *partie linéaire principale* de f) et se note $(\mathrm{d}f)_p$ (ou $\mathrm{T}_p f$).

Exercice 4. L'image d'une courbe (5) de la surface $\mathscr{X}$ par l'application f est une courbe

$$\hat{u}=\hat{u}\,(u\,(t),\ v\,(t)), \quad \hat{v}=\hat{v}\,(u\,(t),\ v\,(t)), \quad t\in I, \tag{25}$$

de la surface $\hat{\mathscr{X}}$. Supposons que pour $t = t_0$ la courbe (5) passe par un point p_0 de $\mathscr{X}$. Montrer que l'image du vecteur tangent à la courbe (5) au point p_0 par la différentielle $(\mathrm{d}f)_{p_0}$ de l'application f est le vecteur tangent à la courbe (25) au point $f\,(p_0)$.

* * *

Une application $f\colon \mathscr{X} \to \hat{\mathscr{X}}$ est un *difféomorphisme* si l'application $\bar{f}$ l'est. (Il est clair que cette définition est intrinsèque.) On sait du cours d'analyse que le jacobien des fonctions (22) définissant le difféomorphisme est non nul:

$$\begin{vmatrix} \dfrac{\partial \hat{u}}{\partial u} & \dfrac{\partial \hat{u}}{\partial v} \\[2ex] \dfrac{\partial \hat{v}}{\partial u} & \dfrac{\partial \hat{v}}{\partial v} \end{vmatrix} \neq 0.$$

Ceci exprime (cf. formules (24)) que *la différentielle* $(\mathrm{d}f)_p$ *du difféomorphisme* f *en tout point* $p \in \mathscr{X}$ *est un isomorphisme entre les espaces vectoriels* $\mathrm{T}_p\mathscr{X}$ *et* $\mathrm{T}_{\hat{p}}\hat{\mathscr{X}}$.

Pour $\hat{U} = U$ les formules $\hat{u} = u$ et $\hat{v} = v$ définissent manifestement un difféomorphisme. On dit de ce difféomorphisme qu'il *opère par égalisation des coordonnées.*

Il est intéressant de noter que *pour tout difféomorphisme* $f: \mathcal{X} \to \hat{\mathcal{X}}$ *et toute paramétrisation* $\gamma: U \to \mathcal{A}$ *de la surface* $\mathcal{X}$ *il existe une paramétrisation* $\gamma^*: U \to \mathcal{A}$ *de la surface* $\hat{\mathcal{X}}$ *telle que dans les paramétrisations* γ *et* γ^* *le difféomorphisme* f *opère par égalisation des coordonnées.* En effet, supposons que $\hat{\gamma}: \hat{U} \to \mathcal{A}$ est une paramétrisation arbitraire de la surface $\hat{\mathcal{X}}$ et que dans les paramétrisations γ et $\hat{\gamma}$ l'application f est représentée par le difféomorphisme $\bar{f}: U \to \hat{U}$. Considérons l'application composée $\gamma^* = \hat{\gamma} \circ \bar{f}$. Etant un difféomorphisme, l'application $\bar{f}$ est une paramétrisation de la surface $\hat{\mathcal{X}}$ équivalente à la paramétrisation $\hat{\gamma}$. Puisque $f \circ \gamma = \hat{\gamma} \circ \bar{f} = \gamma^* \circ \mathrm{id}$, où id est l'application identique, le difféomorphisme f est représenté par l'application id dans les paramétrisations γ et γ^*, i.e. opère par égalisation des coordonnées. □

Nous verrons que cette propriété facilite considérablement parfois l'étude des difféomorphismes.

* * *

On admettra maintenant que l'espace $\mathcal{A}$ qui contient la surface $\mathcal{X}$ est *euclidien.* Tout sous-espace tangent $T_p\mathcal{X}$ hérite alors de la structure euclidienne de cet espace et, de plus, le carré de la longueur d'un vecteur arbitraire (18) de ce sous-espace est défini par la formule

$$\mathrm{d}\mathbf{r}^2 = E\,\mathrm{d}u^2 + 2F\,\mathrm{d}u\,\mathrm{d}v + G\,\mathrm{d}v^2, \tag{26}$$

où

$$E = \mathbf{r}_u^2, \quad F = \mathbf{r}_u\mathbf{r}_v, \quad G = \mathbf{r}_v^2$$

sont les coefficients métriques de la base $\mathbf{r}_u$, $\mathbf{r}_v$.

Définition 5. La forme quadratique (26) des variables $\mathrm{d}u$ et $\mathrm{d}v$ s'appelle *première forme quadratique* de la surface $\mathcal{X}$. On la désigne généralement par le symbole I ou I (dr). Donc, par définition,

$$\mathrm{I} = \mathrm{d}\mathbf{r}^2.$$

Soulignons que les coefficients E, F et G de la première forme quadratique dépendent d'un point $p \in \mathcal{X}$ et en coordonnées u, v sont des fonctions différentiables

$$E = E(u, v), \quad F = F(u, v), \quad G = G(u, v)$$

de u et v.

Remarque 2. L'expression $E\,\mathrm{d}u^2 + 2F\,\mathrm{d}u\,\mathrm{d}v + G\,\mathrm{d}v^2$ est traitée traditionnellement comme une forme quadratique. D'un point de

vue moderne, il convient de la regarder comme l'écriture d'une fonctionnelle quadratique sur l'espace $T_p\mathscr{X}$ prenant la valeur $E\,du^2 + F\,du\,dv + G\,dv^2$ sur le vecteur (18).

La longueur $|\mathbf{r}'(t)|$ du vecteur $\mathbf{r}'(t)$ tangent à la courbe (6) sur la surface $\mathscr{X}$ est définie en vertu de (16) par la formule

$$|\mathbf{r}'(t)| = \sqrt{\mathbf{r}'(t)^2} = \sqrt{(u'(t)\,\mathbf{r}_u + v'(t)\,\mathbf{r}_v)^2} = \\ = \sqrt{Eu'(t)^2 + 2Fu'(t)\,v'(t) + Gv'(t)^2},$$

où E, F et G sont naturellement traités comme des fonctions de t:

$$E = E(u(t),\ v(t)),\quad F = F(u(t),\ v(t)),\quad G = G(u(t),\ v(t)).$$

Pour la longueur s de la courbe (6) on obtient la formule

$$s = \int_a^b \sqrt{Eu'(t)^2 + 2Fu'(t)\,v'(t) + Gv'(t)^2}\,dt, \tag{27}$$

qui peut être mise sous la forme conventionnelle suivante plus facile à retenir

$$s = \int_L \sqrt{E\,du^2 + 2F\,du\,dv + G\,dv^2},$$

où L désigne la courbe (6). La formule

$$ds^2 = E\,du^2 + F\,du\,dv + G\,dv^2 \tag{28}$$

est d'une forme plus conventionnelle encore (elle s'écrit brièvement $ds^2 = \mathrm{I}(d\mathbf{r})$ ou $ds^2 = d\mathbf{r}^2$). Cette relation exprime que *la première forme quadratique* I *définit le carré* ds^2 *de l'élément de longueur.*

[Insistons bien sur le fait que cet énoncé et la formule (28) sont conventionnels. Ils ne servent qu'à écrire la formule (27) sous une forme abrégée.]

On sait du cours d'algèbre linéaire que l'angle θ de deux vecteurs tangents

$$d\mathbf{r} = \mathbf{r}_u\,du + \mathbf{r}_v\,dv \quad \text{et} \quad \delta\mathbf{r} = \mathbf{r}_u\delta u + \mathbf{r}_v\delta v$$

est défini par la formule

$$\cos\theta = \frac{d\mathbf{r}\,\delta\mathbf{r}}{|d\mathbf{r}|\,|\delta\mathbf{r}|},$$

i.e. par la formule

$$\cos\theta = \frac{E\,du\,\delta u + F(du\,\delta v + \delta u\,dv) + G\,dv\,\delta v}{\sqrt{E\,du^2 + 2F\,du\,dv + G\,dv^2}\,\sqrt{E\,\delta u^2 + 2F\,\delta u\,\delta v + G\,\delta v^2}}, \tag{29}$$

que l'on conviendra d'écrire sous la forme mnémonique suivante:

$$\cos\theta = \frac{\mathrm{I}(d,\ \delta)}{\sqrt{\mathrm{I}(d)}\,\sqrt{\mathrm{I}(\delta)}}.$$

On appelle *angle des courbes*

$$\mathbf{r} = \mathbf{r}(u(t),\ v(t)) \quad \text{et} \quad \mathbf{r}_1 = \mathbf{r}_1(u_1(t),\ v_1(t))$$

d'une surface $\mathcal{X}$, passant par un même point p_0 de $\mathcal{X}$ pour $t = t_0$, l'angle de leurs vecteurs tangents

$$\mathbf{r}'(t_0) = u'(t_0)\,\mathbf{r}_{u_0} + v'(t_0)\,\mathbf{r}_{v_0} \quad \text{et} \quad \mathbf{r}_1'(t_0) = u_1'(t_0)\,\mathbf{r}_{1u_0} + v_1'(t_0)\,\mathbf{r}_{1v_0}$$

en p_0. On a

$$\cos\theta = \frac{\mathbf{r}'(t_0)\,\mathbf{r}_1'(t_0)}{|\mathbf{r}'(t_0)|\cdot|\mathbf{r}_1'(t_0)|},$$

où

$$\mathbf{r}'(t_0)\,\mathbf{r}_1'(t_0) =$$
$$= E_0 u'(t_0)\,u_1'(t_0) + F_0(u'(t_0)\,v_1'(t_0) + u_1'(t_0)\,v'(t_0)) + G_0 v'(t_0)\,v_1'(t_0),$$
$$|\mathbf{r}'(t_0)| = \sqrt{E_0 u'(t_0)^2 + 2F_0 u'(t_0)\,v'(t_0) + G_0 v'(t_0)^2},$$
$$|\mathbf{r}_1'(t_0)| = \sqrt{E_0 u_1'(t_0)^2 + 2F_0 u_1'(t_0)\,v_1'(t_0) + G_0 v_1'(t_0)^2},$$

et

$$E_0 = E(u_0,\ v_0) = E(u(t_0),\ v(t_0)),$$
$$F_0 = F(u_0,\ v_0) = F(u(t_0),\ v(t_0)),$$
$$G_0 = G(u_0,\ v_0) = G(u(t_0),\ v(t_0))$$

sont les valeurs des coefficients de la première forme quadratique au point p_0.

En particulier, le cosinus de l'angle des lignes de coordonnées $u = \text{const}$ et $v = \text{const}$ est défini par la formule

$$\cos\theta = \frac{F}{\sqrt{E}\,\sqrt{G}},$$

d'où il s'ensuit que *les lignes de coordonnées* $u = \text{const}$ et $v = \text{const}$ *sont orthogonales si et seulement si* $F = 0$.

Nous pouvons donc calculer les longueurs des courbes d'une surface et les angles de ces courbes en sachant seulement la première forme quadratique de cette surface, i.e. les structures euclidiennes des plans tangents.

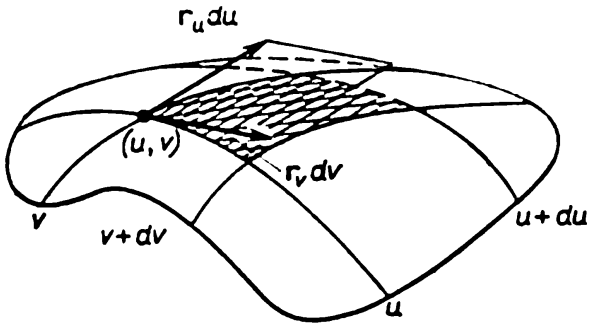
Du cours d'analyse on sait que l'*aire* d'une portion D de la surface $\mathcal{X}$ est définie comme la limite des aires des polygones approchant D. Cette limite (si elle existe) s'exprime par l'intégrale

$$\iint_{\Gamma} \sqrt{\Gamma(\mathbf{r}_u,\ \mathbf{r}_v)}\,du\,dv,$$

où

$$\Gamma = (\mathbf{r}_u, \mathbf{r}_v) = \begin{vmatrix} \mathbf{r}_u^2 & \mathbf{r}_u\mathbf{r}_v \\ \mathbf{r}_u\mathbf{r}_v & \mathbf{r}_v^2 \end{vmatrix} = \begin{vmatrix} E & F \\ F & G \end{vmatrix} = EG - F^2$$

est le *déterminant de Gram* des vecteurs $\mathbf{r}_u$ et $\mathbf{r}_v$. [La justification intuitive de cette formule est la suivante: on voit sur la figure que l'aire d'un parallélogramme curviligne infiniment petit du réseau de coordonnées, de sommet (u, v) et de côtés du et dv est approximativement égale à celle du parallélogramme du plan tangent, tendu sur $\mathbf{r}_u du$ et $\mathbf{r}_v dv$. Par ailleurs, il est aisé de prouver, — cf. lemme 1, leçon II.26, — que cette aire est égale à $\sqrt{\Gamma(\mathbf{r}_u, \mathbf{r}_v)}\, du\, dv$.]



En analyse, on exprime usuellement ce fait en disant que *l'élément d'aire d'une surface est égal à* $\sqrt{EG - F^2}\, du\, dv$.

Ainsi, la première forme quadratique d'une surface nous permet de calculer toute aire de cette surface.

* * *

Pour abréger on appellera *isométrie* un isomorphisme linéaire des espaces euclidiens. Par définition (cf. définition 7, leçon II.14), un isomorphisme linéaire $\varphi: \mathscr{V} \to \mathscr{V}_1$ est une isométrie si

$$\mathbf{x}\mathbf{y} = \varphi\mathbf{x} \cdot \varphi\mathbf{y} \tag{30}$$

pour tous vecteurs $\mathbf{x}, \mathbf{y} \in \mathscr{V}$. Ceci étant, il suffit d'exiger la réalisation de la condition (30) uniquement pour $\mathbf{x} = \mathbf{y}$, i.e. *une application linéaire* $\varphi: \mathscr{V} \to \mathscr{V}_1$ *telle que*

$$\mathbf{x}^2 = (\varphi\mathbf{x})^2 \quad \textit{pour tout vecteur } \mathbf{x} \in \mathscr{V} \tag{31}$$

est une isométrie. (Pour prouver ceci il suffit d'appliquer la relation (31) au vecteur $\mathbf{x} + \mathbf{y}$.)

En coordonnées, l'application φ est définie par les formules

$$y^i = a^i_j x^j,$$

et les carrés scalaires $\mathbf{x}^2$, $\mathbf{x} \in \mathscr{V}$, et $\mathbf{y}^2$, $\mathbf{y} \in \mathscr{V}_1$, par les formes quadratiques

$$\mathbf{x}^2 = g_{ij}x^i x^j, \quad \mathbf{y}^2 = h_{ij}y^i y^j.$$

Dans ces notations, l'égalité (31) exprime que la substitution des expressions $y^i = a^i_j x^j$ dans la forme $h_{ij}y^iy^j$ doit nous donner la forme $g_{ij}x^ix^j$, i.e. on a les égalités

$$g_{ij} = h_{kl}a^k_i a^l_j.$$

En particulier, pour les espaces euclidiens à deux dimensions $\mathrm{T}_p\mathscr{X}$, $\mathrm{T}_{\hat{p}}\hat{\mathscr{X}}$ et la différentielle $(\mathrm{d}f)_p : \mathrm{T}_p\mathscr{X} \to \mathrm{T}_{\hat{p}}\hat{\mathscr{X}}$, $\hat{p} = f(p)$, du difféomorphisme $f : \mathscr{X} \to \hat{\mathscr{X}}$, la réalisation de la condition (31) en tout point $p \in \mathscr{X}$ exprime que la substitution des expressions (24) de $\mathrm{d}\hat{u}$ et $\mathrm{d}\hat{v}$ dans la première forme quadratique $\hat{E}\,\mathrm{d}\hat{u}^2 + 2\hat{F}\,\mathrm{d}\hat{u}\,\mathrm{d}\hat{v} + \hat{G}\,\mathrm{d}\hat{v}^2$ de la surface $\hat{\mathscr{X}}$ (et la substitution des expressions (22) de $\hat{u}$ et $\hat{v}$ dans ses coefficients $\hat{E}$, $\hat{F}$ et $\hat{G}$) nous donne la première forme quadratique $E\,\mathrm{d}u^2 + 2F\,\mathrm{d}u\,\mathrm{d}v + G\,\mathrm{d}v^2$ de la surface $\mathscr{X}$, i.e. on a identiquement en u et v

$$\begin{aligned}
&E(u, v) = \\
&\quad = \hat{E}(\hat{u}, \hat{v})\left(\frac{\partial\hat{u}}{\partial u}\right)^2 + 2\hat{F}(\hat{u}, \hat{v})\frac{\partial\hat{u}}{\partial u}\frac{\partial\hat{v}}{\partial u} + \hat{G}(\hat{u}, \hat{v})\left(\frac{\partial\hat{v}}{\partial u}\right)^2, \\
&F(u, v) = \\
&\quad = \hat{E}(\hat{u}, \hat{v})\frac{\partial\hat{u}}{\partial u}\frac{\partial\hat{u}}{\partial v} + \hat{F}(\hat{u}, \hat{v})\left(\frac{\partial\hat{u}}{\partial u}\frac{\partial\hat{v}}{\partial v} + \frac{\partial\hat{u}}{\partial v}\frac{\partial\hat{v}}{\partial u}\right) + \\
&\qquad\qquad + \hat{G}(\hat{u}, \hat{v})\frac{\partial\hat{v}}{\partial u}\frac{\partial\hat{v}}{\partial v}, \\
&G(u, v) = \\
&\quad = \hat{E}(\hat{u}, \hat{v})\left(\frac{\partial\hat{u}}{\partial v}\right)^2 + 2\hat{F}(\hat{u}, \hat{v})\frac{\partial\hat{u}}{\partial v}\frac{\partial\hat{v}}{\partial v} + \hat{G}(\hat{u}, \hat{v})\left(\frac{\partial\hat{v}}{\partial v}\right)^2.
\end{aligned} \tag{32}$$

Donc, *les applications linéaires*

$$(\mathrm{d}f)_p : \mathrm{T}_p\mathscr{X} \to \mathrm{T}_{\hat{p}}\hat{\mathscr{X}}, \quad \hat{p} = f(p),$$

où $f : \mathscr{X} \to \hat{\mathscr{X}}$ est un difféomorphisme, sont des isométries pour tout point $p \in \mathscr{X}$ si et seulement si les égalités (32) *ont lieu identiquement en u et v.*

Pour un difféomorphisme $f : \mathscr{X} \to \hat{\mathscr{X}}$ opérant par égalisation des coordonnées, cette condition exprime que *les premières formes quadratiques des surfaces $\mathscr{X}$ et $\hat{\mathscr{X}}$ sont confondues* dans les coordonnées envisagées (ou plus exactement qu'elles ne diffèrent que par les notations des variables).

Définition 6. On dit qu'un difféomorphisme $f: \mathscr{X} \to \hat{\mathscr{X}}$ est une *application isométrique* (ou simplement une *isométrie*) de la surface $\mathscr{X}$ sur la surface $\hat{\mathscr{X}}$ si elle envoie toute courbe

$$u = u(t), \quad v = v(t), \quad a \leqslant t \leqslant b, \tag{33}$$

de la surface $\mathscr{X}$ dans une courbe de même longueur

$$\hat{u} = \hat{u}(t), \quad \hat{v} = \hat{v}(t), \quad a \leqslant t \leqslant b,$$

de la surface $\hat{\mathscr{X}}$, où $\hat{u}(t) = \hat{u}(u(t), v(t))$ et $\hat{v}(t) = \hat{v}(u(t), v(t))$, i.e. si

$$\int_a^b \sqrt{E(t)\,u'(t)^2 + 2F(t)\,u'(t)\,v'(t) + G(t)\,v'(t)^2}\,dt =$$

$$= \int_a^b \sqrt{\hat{E}(t)\,\hat{u}'(t)^2 + 2\hat{F}(t)\,\hat{u}'(t)\,\hat{v}'(t) + \hat{G}(t)\,\hat{v}'(t)^2}\,dt,$$

où

$$E(t) = E(u(t), v(t)), \quad F(t) = F(u(t), v(t)),$$
$$G(t) = G(u(t), v(t))$$

et de façon analogue

$$\hat{E}(t) = \hat{E}(\hat{u}(t), \hat{v}(t)), \quad \hat{F}(t) = \hat{F}(\hat{u}(t), \hat{v}(t)),$$
$$\hat{G}(t) = \hat{G}(\hat{u}(t), \hat{v}(t)).$$

Les surfaces pour lesquelles il existe au moins une application isométrique $\mathscr{X} \to \hat{\mathscr{X}}$ sont dites *isométriques*.

Proposition 1. *Un difféomorphisme $f: \mathscr{X} \to \hat{\mathscr{X}}$ est une isométrie si et seulement si sa différentielle*

$$(df)_p : \mathbf{T}_p\mathscr{X} \to \mathbf{T}_{\hat{p}}\hat{\mathscr{X}}, \quad \hat{p} = f(p),$$

l'est pour tout point $p \in \mathscr{X}$.

D é m o n s t r a t i o n. D'après les remarques faites plus haut, on peut, sans perdre en généralité, admettre que le difféomorphisme f opère par égalisation des coordonnées. L'image de toute courbe (33) de la surface $\mathscr{X}$ par le difféomorphisme f sera alors une courbe de $\hat{\mathscr{X}}$ définie par les mêmes équations paramétriques que (33), et affirmer que la différentielle $(df)_p$ du difféomorphisme f est une isométrie en tout point $p \in \mathscr{X}$ revient à affirmer que les premières formes quadratiques des surfaces $\mathscr{X}$ et $\hat{\mathscr{X}}$ sont les mêmes aux notations près

des variables. Donc, les formules (27) seront les mêmes pour les deux courbes et, par conséquent, les longueurs de ces courbes seront égales. Donc, f est une isométrie.

Supposons réciproquement que le difféomorphisme $f\colon \mathscr{X} \to \hat{\mathscr{X}}$ qui opère par égalisation des coordonnées est une isométrie. Ceci signifie que pour toutes fonctions différentiables $u = u(t)$, $v = v(t)$, $a \leqslant t \leqslant b$, telles que $(u(t), v(t)) \in U$, $a \leqslant t \leqslant b$, on a l'égalité

$$\int_a^b \sqrt{E(t)\,u'(t)^2 + 2F(t)\,u'(t)\,v'(t) + G(t)\,v'(t)^2}\,\mathrm{d}t =$$

$$= \int_a^b \sqrt{\hat{E}(t)\,u'(t)^2 + 2\hat{F}(t)\,u'(t)\,v'(t) + \hat{G}(t)\,v'(t)^2}\,\mathrm{d}t.$$

En dérivant cette identité par rapport à b (et en remplaçant b par t) on obtient, après élévation au carré, l'identité

$$E(t)\,u'(t)^2 + 2F(t)\,u'(t)\,v'(t) + G(t)\,v'(t)^2 =$$

$$= \hat{E}(t)\,u'(t)^2 + 2\hat{F}(t)\,u'(t)\,v'(t) + \hat{G}(t)\,v'(t)^2.$$

Cette identité doit en particulier avoir lieu pour toute fonction affine de la forme

$$u(t) = u_0 + \alpha t, \quad v(t) = v_0 + \beta t, \quad |t| < \varepsilon,$$

où (u_0, v_0) est un point arbitraire d'un domaine U, α et β des nombres arbitraires et $\varepsilon > 0$ un nombre assez petit. Mais dans ce cas cette identité devient par la substitution $t = 0$

$$E_0\alpha^2 + 2F_0\alpha\beta + G_0\beta^2 = \hat{E}_0\alpha^2 + 2\hat{F}_0\alpha\beta + \hat{G}_0\beta^2,$$

où $E_0, \ldots, \hat{G}_0$ sont les valeurs des fonctions $E, \ldots, \hat{G}$ au point (u_0, v_0), et donc n'a lieu, en raison du choix arbitraire de α et β, que si $E_0 = \hat{E}_0$, $F_0 = \hat{F}_0$, $G_0 = \hat{G}_0$, i.e. si $E = \hat{E}$, $F = \hat{F}$ et $G = \hat{G}$ partout dans U. Donc, les applications linéaires $(\mathrm{d}f)_p$ sont des isométries. □

Corollaire 1. *Les courbes homologues de surfaces isométriques se coupent sous le même angle et les domaines homologues sont de même aire.* □

Corollaire 2. *Deux surfaces sont isométriques si et seulement si elles peuvent être munies de coordonnées locales dans lesquelles leurs premières formes quadratiques sont confondues.* □

Ce critère est certes au plus haut point inefficace (comment en effet deviner l'existence de ces coordonnées?). L'objectif que l'on se fixe (et qui sera atteint à la leçon 5) est de le rendre efficace, mais le chemin sera assez long.

Si l'on considère une surface flexible et inextensible et qu'on la torde arbitrairement, on ne modifiera pas les longueurs de ses courbes et on obtiendra donc une surface isométrique. En partant de cette constatation, les fondateurs de la théorie des surfaces du XIXe siècle ont appelé les isométries *déformations*. Cette terminologie a subsisté partiellement jusqu'à nos jours, mais actuellement les déformations sont comprises généralement sous un sens plus étroit, plus exactement comme des isométries qui peuvent être reliées à l'application identique par une famille continue d'isométries.

Longtemps les mathématiciens étaient convaincus que localement, i.e. dans un voisinage assez petit d'un point arbitraire, toute isométrie est une déformation dans ce sens. Mais tout dernièrement N. Efimov a réfuté ce point de vue en construisant un contre-exemple *ad hoc*.

Supposons qu'une surface est peuplée d'êtres doués de raison, capables de mesurer longueurs, aires et angles, mais incapables d'accéder à l'espace ambiant. Leur géométrie restera inchangée par toute déformation de la surface et ils ne remarqueront même pas cette déformation. Pour cette raison on dit des surfaces isométriques qu'elles ont la même *géométrie intrinsèque*.

* * *

Citons un important exemple de construction géométrique intrinsèque.

En vertu de résultats généraux d'algèbre linéaire (cf. leçon II.5), l'espace dual $\mathcal{V}^*$ de tout espace euclidien $\mathcal{V}$ muni d'un tenseur métrique g_{ij} est un espace euclidien muni d'un tenseur métrique g^{ij}, dont les composantes forment une matrice $\| g^{ij} \|$ inverse de la matrice $\| g_{ij} \|$ des composantes du tenseur g_{ij}. Donc, pour tout covecteur $\xi = (\xi_1, \ldots, \xi_n)$ de $\mathcal{V}^*$ est définie la longueur $| \xi |$ dont le carré est donné par la formule

$$| \xi |^2 = g^{ij}\xi_i\xi_j.$$

Dans notre cas, pour l'espace euclidien à deux dimensions $T_p\mathcal{X}$ la matrice $\| g_{ij} \|$ est de la forme

$$\begin{Vmatrix} E & F \\ F & G \end{Vmatrix},$$

donc la matrice $\| g^{ij} \|$ est de la forme

$$\frac{1}{EG-F^2}\begin{Vmatrix} G & -F \\ -F & E \end{Vmatrix}.$$

Pour la longueur $|\xi|$ d'un covecteur arbitraire ξ de $T_p\mathscr{X}$ on a donc

$$|\xi|^2 = \frac{G\xi^2 - 2F\xi\eta + E\eta^2}{EG - F^2},$$

où ξ et η sont les coordonnées de ce covecteur.

Un exemple de covecteur ξ est le gradient $\left(\frac{\partial\varphi}{\partial u}, \frac{\partial\varphi}{\partial v}\right)$ d'une fonction différentiable $\varphi = \varphi(u, v)$ sur $\mathscr{X}$. Le carré de la longueur de ce covecteur s'appelle *premier paramètre différentiel de Beltrami* de la fonction φ et se note $\Delta_1\varphi$. Donc, par définition,

$$\Delta_1\varphi = \frac{G\varphi_u^2 - 2F\varphi_u\varphi_v + E\varphi_v^2}{EG - F^2}.$$

Cette construction est intrinsèque, i.e. ne dépend pas du choix de la paramétrisation, et pour toute isométrie $f: \mathscr{X} \to \hat{\mathscr{X}}$ on a la formule

$$\Delta_1(\varphi \circ f) = \Delta_1\varphi \circ f \tag{34}$$

(vérifiez-le !) Dans ce sens, elle appartient à la géométrie intrinsèque de la surface $\mathscr{X}$.

* * *

Considérons en conclusion quelques exemples de calcul de la première forme quadratique de surfaces gauches. En principe, ces surfaces ne seront pas élémentaires, mais on pourra facilement les y ramener par découpage et restriction des domaines de définition des paramétrisations.

Exemple 4. L'équation paramétrique du plan Oxy rapporté aux coordonnées $u = x$ et $v = y$ est $\mathbf{r} = u\mathbf{i} + v\mathbf{j}$. Donc, $\mathbf{r}_u = \mathbf{i}$, $\mathbf{r}_v = \mathbf{j}$ et, par suite, $E = 1$, $F = 0$, $G = 1$, i.e. pour le plan,

$$\mathrm{I} = du^2 + dv^2. \tag{35}$$

(Un résultat qui se devine facilement sans le moindre calcul.)

Il va de soi que tout sous-ensemble ouvert du plan (traité comme une surface gauche) possède la même première forme quadratique.

Exemple 5. Pour le cylindre circulaire

$$\mathbf{r} = R\cos u\cdot\mathbf{i} + R\sin u\cdot\mathbf{j} + v\cdot\mathbf{k}$$

on a $\mathbf{r}_u = -R\sin u\cdot\mathbf{i} + R\cos u\cdot\mathbf{j}$ et $\mathbf{r}_v = \mathbf{k}$. Donc,

$$E = \mathbf{r}_u^2 = R^2, \quad F = \mathbf{r}_u\mathbf{r}_v = 0, \quad G = \mathbf{r}_v^2 = 1,$$

i.e. pour le cylindre,

$$\mathrm{I} = R^2 du^2 + dv^2.$$

En introduisant la nouvelle coordonnée $u_1 = Ru$ (et en désignant u_1 de nouveau par u), on ramène cette expression à la forme (35).

Il existe donc des coordonnées dans lesquelles les premières formes quadratiques du plan et du cylindre sont confondues! Mais cela ne signifie encore pas que le plan et le cylindre (bien sûr découpé, cf. exemple 1 ci-dessus) sont isométriques, puisque les coordonnées du cylindre ne parcourent qu'une bande de $\mathbb{R}^2$, donc le cylindre découpé est isométrique à une partie seulement du plan. On exprimera ce fait en disant que le cylindre et le plan sont *localement isométriques*.

En déformant petit à petit un cylindre découpé, on l'applique isométriquement sur une bande de plan.

Exemple 6. Pour la surface de révolution

$$\mathbf{r} = x(v)\cos u\cdot\mathbf{i} + x(v)\sin u\cdot\mathbf{j} + z(v)\cdot\mathbf{k}$$

on a

$$\mathbf{r}_u = -x(v)\sin u\cdot\mathbf{i} + x(v)\cos u\cdot\mathbf{j},$$
$$\mathbf{r}_v = x'(v)\cos u\cdot\mathbf{i} + x'(v)\sin u\cdot\mathbf{j} + z'(v)\,\mathbf{k}.$$

Donc,

$$E = x(v)^2\sin^2 u + x(v)^2\cos^2 u = x(v)^2,$$
$$F = -x(v)\sin u\cdot x'(v)\cos u + x(v)\cos u\cdot x'(v)\sin u = 0,$$
$$G = x'(v)^2\cos^2 u + x'(v)^2\sin^2 u + z'(v)^2 = x'(v)^2 + z'(v)^2,$$

de sorte que

$$\mathrm{I} = x(v)^2\,\mathrm{d}u^2 + (x'(v)^2 + z'(v)^2)\,\mathrm{d}v^2.$$

Il est évident que les méridiens et les parallèles de toute surface de révolution sont orthogonaux. Donc, on aurait pu établir que $F = 0$ sans le moindre calcul.

Lorsque le profil $x = x(v)$, $z = z(v)$ d'une surface de révolution est paramétré par un paramètre naturel $v = s$ (et donc que $x'(v)^2 + z'(v)^2 = 1$), la forme I s'écrit tout simplement

$$\mathrm{I} = x(v)^2\,\mathrm{d}u^2 + \mathrm{d}v^2.$$

Il s'avère en particulier que *la première forme quadratique d'une sphère de rayon* 1 *est*

$$\mathrm{I} = \cos^2 v\,\mathrm{d}u^2 + \mathrm{d}v^2. \tag{36}$$

Les cartographes savent qu'aucune portion, même petite, de sphère ne peut être appliquée sur un plan. Cela signifie qu'aucun changement de coordonnées ne peut transformer la forme (36) en la forme (35). Mais comment le prouver? La réponse sera donnée à la leçon 5.

Exemple 7. La *chaînette* est la configuration d'un fil pesant homogène dont les extrémités sont fixées, et la *caténoïde*, la surface de révolution engendrée par la chaînette.

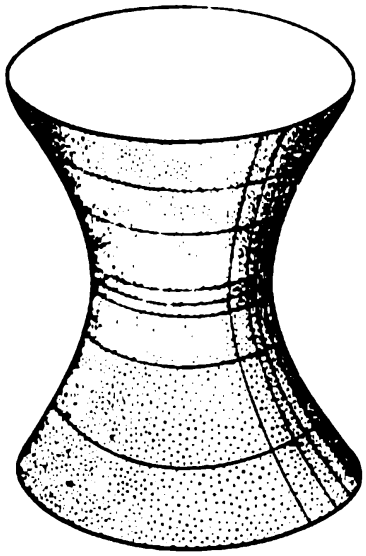
En mécanique (statique) on démontre que la chaînette est le graphique du cosinus hyperbolique. Donc, pour la caténoïde, $x(v) = \operatorname{ch} v$, $z(v) = v$ et, par suite,

$$x(v)^2 = \operatorname{ch}^2 v \text{ et } x'(v)^2 + z'(v)^2 = \operatorname{sh}^2 v + 1 = \operatorname{ch}^2 v.$$

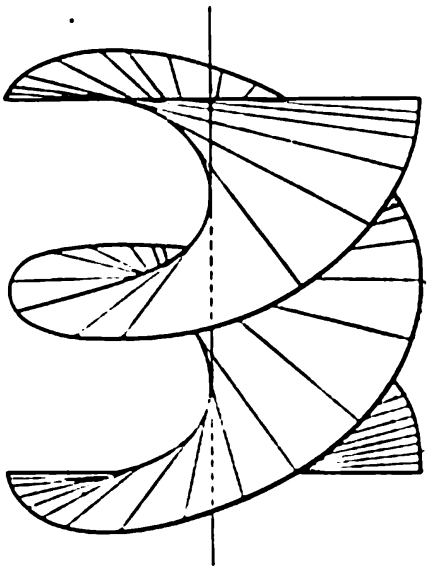
Donc, pour la caténoïde,

$$\mathrm{I} = \operatorname{ch}^2 v\,(du^2 + dv^2). \tag{37}$$

Exemple 8. Supposons qu'une droite perpendiculaire à l'axe Oz tourne uniformément autour de cet axe et se déplace simultané-



Caténoïde



Hélicoïde

ment vers le haut (d'une hauteur proportionnelle à l'angle de rotation). La surface réglée balayée par cette droite s'appelle *hélicoïde*.

Si v est un paramètre sur la droite, et u, l'angle de rotation, l'hélicoïde sera défini par l'équation

$$\mathbf{r} = v \cos u \cdot \mathbf{i} + v \sin u \cdot \mathbf{j} + u \cdot \mathbf{k}.$$

Donc,

$$\mathbf{r}_u = -v \sin u \cdot \mathbf{i} + v \cos u \cdot \mathbf{j} + \mathbf{k},$$
$$\mathbf{r}_v = \cos u \cdot \mathbf{i} + \sin u \cdot \mathbf{j}$$

et, par suite,

$$E = 1 + v^2, \quad F = 0, \quad G = 1.$$

Donc, pour l'hélicoïde,

$$\mathrm{I} = (1 + v^2)\,du^2 + dv^2.$$

Transformons cette forme en introduisant de nouvelles coordonnées u_1, v_1 liées aux coordonnées u, v par les formules

$$u = u_1, \quad v = \operatorname{sh} v_1.$$

Alors

$$1 + v^2 = 1 + \operatorname{sh}^2 v_1 = \operatorname{ch}^2 v_1,$$
$$du = du_1, \quad dv = \operatorname{ch} v_1 \, dv_1,$$

donc (on omet les indices des nouvelles coordonnées)

$$\mathrm{I} = \operatorname{ch}^2 v \, (du^2 + dv^2),$$

ce qui est confondu avec la forme (37).

Ceci prouve que *la caténoïde et l'hélicoïde sont localement isométriques* (plus exactement, la caténoïde découpée le long d'un méridien est isométrique à la partie $0 < u < 2\pi$ de l'hélicoïde), et de plus il existe une isométrie transformant les méridiens de la caténoïde en les génératrices rectilignes de l'hélicoïde.

Exemple 9. Pour une surface réglée arbitraire

$$\mathbf{r} = \boldsymbol{\rho}(u) + v\mathbf{a}(u), \tag{38}$$

où (cf. exemple 3 plus haut) $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$ est une courbe régulière de paramètre naturel, et $\mathbf{a}(u)$ une fonction vectorielle telle que $|\mathbf{a}(u)| = 1$ pour tout u, on obtient, en désignant par un point la dérivation par rapport à u,

$$\mathbf{r}_u = \dot{\boldsymbol{\rho}} + v\dot{\mathbf{a}}, \quad \mathbf{r}_v = \mathbf{a}.$$

Puisque $\dot{\boldsymbol{\rho}}^2 = 1$ et $\mathbf{a}^2 = 1$, $\mathbf{a}\dot{\mathbf{a}} = 0$, il vient

$$E = 1 + 2v\dot{\boldsymbol{\rho}}\dot{\mathbf{a}} + v^2\dot{\mathbf{a}}^2, \quad F = \dot{\boldsymbol{\rho}}\mathbf{a}, \quad G = 1.$$

Si en particulier $\mathbf{a} = \dot{\boldsymbol{\rho}}$ (la surface des tangentes), alors $\dot{\boldsymbol{\rho}}\mathbf{a} = \mathbf{a}^2 = 1$ (i.e $F = 1$) et $\dot{\boldsymbol{\rho}}\dot{\mathbf{a}} = 0$, $\dot{\mathbf{a}}^2 = k^2$, où k est la courbure de la courbe $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$ (i.e. $E = 1 + k^2v^2$). Donc, pour la surface des tangentes,

$$\mathrm{I} = (1 + k^2v^2)\, du^2 + 2du\, dv + dv^2. \tag{39}$$

Si $\mathbf{a}(u)$ est le vecteur binormal à la courbe $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$, alors $\dot{\boldsymbol{\rho}}\mathbf{a} = 0$, $\dot{\boldsymbol{\rho}}\dot{\mathbf{a}} = 0$ et $\dot{\mathbf{a}}^2 = \varkappa^2$, où $\varkappa$ est la torsion de la courbe $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$. Donc, pour la surface des binormales,

$$\mathrm{I} = (1 + \varkappa^2v^2)\, du^2 + dv^2.$$

On voit donc que la première forme quadratique de la surface des tangentes ne dépend que de la courbure de la courbe envisagée et celle de la surface des binormales, que de la torsion.

* * *

S'agissant de la surface des tangentes, on en déduit que *toute surface des tangentes est localement isométrique à un plan.* En effet, considérons une courbe plane de même courbure $k = k(u)$ (l'existence d'une telle courbe est assurée par le théorème général 1 de la leçon 2). La première forme quadratique de la surface des tangentes de cette courbe sera la forme (39). Mais d'autre part il est clair que la surface des tangentes d'une courbe plane est — en dehors de ses points singuliers — un domaine du plan. Donc, il existe un changement de coordonnées transformant la première forme quadratique $dx^2 + dy^2$ du plan en la forme (39). (Ce changement de coordonnées est de la forme

$$x = x(u) + x'(u)\,v, \qquad y = y(u) + y'(u)\,v,$$

où $x(u)$ et $y(u)$ sont des fonctions telles que $x'(u)^2 + y'(u)^2 = 1$ et $x''(u)^2 + y''(u)^2 = k(u)^2$.) □

On peut réaliser cette isométrie en déformant continûment la courbe $\rho = \rho(u)$ en une courbe plane.

Pour cette raison, les surfaces des tangentes s'appellent *surfaces développables* (sous-entendu, sur le plan).

Lorsque $\mathbf{a}(u) = \rho(u)$, la surface (38) est un cône de sommet en l'origine des coordonnées (et la courbe $\rho = \rho(u)$ est son intersection avec la sphère unité $|\rho| = 1$). Dans ce cas, on a

$$\dot{\rho}\dot{\mathbf{a}} = \dot{\rho}^2 = 1, \quad \dot{\mathbf{a}}^2 = 1, \quad \dot{\rho}\mathbf{a} = 0,$$

d'où l'on déduit l'expression de la première forme quadratique

$$\mathrm{I} = (1 + v)^2\,du^2 + dv^2.$$

Le changement de coordonnées évident $(u, v) \mapsto (u, 1 + v)$ ramène cette forme à la forme plus simple

$$\mathrm{I} = v^2\,du^2 + dv^2. \tag{40}$$

Introduisons maintenant les nouvelles coordonnées

$$x = v\cos u, \qquad y = v\sin u.$$

Alors

$$dx = -v\sin u\,du + \cos u\,dv,$$
$$dy = v\cos u\,du + \sin u\,dv,$$

donc

$$dx^2 + dy^2 = v^2\,du^2 + dv^2.$$

Ceci prouve que *le cône est aussi localement isométrique au plan* (plus exactement, chaque nappe du cône, découpée le long d'une génératrice, est isométrique à un secteur plan ; un fait qui est géo-

métriquement évident). Pour cette raison, les cônes sont des surfaces développables.

Remarquons que la forme (10) n'est autre que la première forme quadratique du plan rapporté aux coordonnées polaires $r = v$ et $\varphi = u$.

Enfin, si le vecteur $\mathbf{a}(u)$ est constant (donc $\dot{\mathbf{a}} = 0$), la surface (38) est un cylindre. Sans perdre en généralité on peut admettre que sa directrice $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$ est une courbe située dans un plan orthogonal au vecteur $\mathbf{a}$ (donc, $\boldsymbol{\rho}\mathbf{a} = 0$ et $\dot{\boldsymbol{\rho}}\mathbf{a} = 0$). Par conséquent, comme pour le cylindre droit (exemple 3),

$$\mathrm{I} = \mathrm{d}u^2 + \mathrm{d}v^2.$$

Pour cette raison, les cylindres font aussi partie des surfaces développables.

A la leçon 4 on montrera que, parmi les surfaces réglées, seules les surfaces développables (i.e. les cylindres, les cônes et les surfaces des tangentes) sont localement isométriques au plan, et à la leçon 5, que les surfaces développables sont les seules surfaces gauches localement isométriques à un plan.

LEÇON 4

Vecteur normal. — La surface, graphique d'une fonction. — Sections normales. — Deuxième forme quadratique d'une surface. — Indicatrice de Dupin. — Courbures principales totale et moyenne. — Deuxième forme quadratique d'un graphique. — Surfaces réglées de courbure nulle. — Surfaces de révolution.

A cette leçon, on se penche plus attentivement sur les surfaces dans un espace euclidien orienté à trois dimensions. En général, les surfaces seront étudiées localement, i.e. dans des voisinages assez petits de leurs points. Donc, on peut, sans perdre en généralité, admettre que ces surfaces sont élémentaires.

* * *

Soient $\mathscr{X}$ une surface élémentaire dans un espace euclidien orienté $\mathscr{A}$ à trois dimensions et $\mathbf{r} = \mathbf{r}(u, v)$ une paramétrisation de $\mathscr{X}$. En tout point $p \in \mathscr{X}$ il existe alors un seul vecteur unitaire $\mathbf{n}$ perpendiculaire au plan tangent et formant une base directe avec les vecteurs $\mathbf{r}_u$ et $\mathbf{r}_v$. Ce vecteur est défini par la formule

$$\mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|}$$

et s'appelle *vecteur normal* à la surface $\mathscr{X}$ en p (et la base $\mathbf{r}_u$, $\mathbf{r}_v$, $\mathbf{n}$ *base mobile* de la surface). Cf. fig. de la page 48.

Par définition, le carré de la longueur $|\mathbf{r}_u \times \mathbf{r}_v|^2$ du produit vectoriel est égal à l'aire du parallélogramme tendu sur les vecteurs $\mathbf{r}_u$ et $\mathbf{r}_v$, i.e. (cf. leçon 3) au déterminant de Gram $\Gamma(\mathbf{r}_u, \mathbf{r}_v) = EG - F^2$ des vecteurs $\mathbf{r}_u$ et $\mathbf{r}_v$. Donc,

$$\mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{\sqrt{EG - F^2}}.$$

Le vecteur $\mathbf{n}$ est généralement calculé à l'aide de cette formule.

* * *

Rapportons $\mathscr{X}$ à des coordonnées rectangulaires x, y, z dont l'origine O est en p et dont l'axe Oz est dirigé par le vecteur $\mathbf{n}$ et considérons la matrice jacobienne

$$\begin{Vmatrix} x_u & y_u & z_u \\ x_v & y_v & z_v \end{Vmatrix}$$

dont les lignes sont les coordonnées des vecteurs $\mathbf{r}_u$ et $\mathbf{r}_v$. La condition imposée à l'axe Oz exprime en particulier que

$$\begin{vmatrix} x_u & y_u \\ x_v & y_v \end{vmatrix} = 1.$$

Donc, en vertu du classique théorème de l'application réciproque, les coordonnées x et y peuvent être exprimées en fonction de u et v dans un voisinage du point (0, 0). En portant ces expressions dans la fonction vectorielle $\mathbf{r} = \mathbf{r}(u, v)$, on obtient une paramétrisation de la surface $\mathscr{X}$ (ou plus exactement d'une portion de $\mathscr{X}$ contenant le point p):

$$\begin{aligned} x &= x, \\ y &= y, \\ z &= z(x, y). \end{aligned}$$

Par définition, cela exprime que *localement* (au voisinage du point p) *la surface $\mathscr{X}$ est le graphique de la fonction $z = z(x, y)$.*

* * *

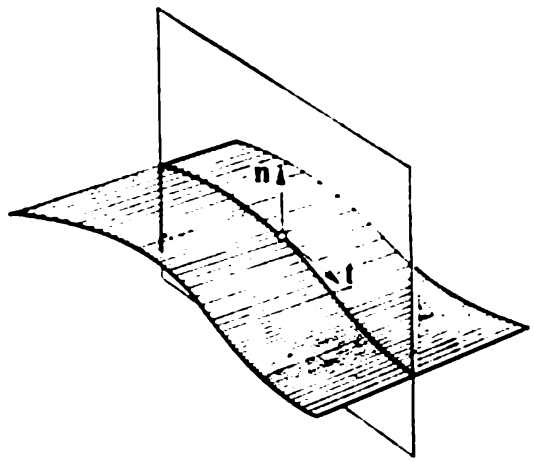
Chaque plan Π contenant l'axe Oz du système de coordonnées envisagé s'appelle *plan normal* à la surface $\mathscr{X}$ au point p et son intersection $\Pi \cap \mathscr{X}$ avec la surface $\mathscr{X}$, *section normale* de $\mathscr{X}$ en p. Tout bivecteur directeur d'un plan normal Π est de la forme $\mathbf{t} \wedge \mathbf{n}$, où $\mathbf{t}$ est un vecteur non nul tangent à $\mathscr{X}$ en p défini à une constante multiplicative près. Le vecteur $\mathbf{t}$ définit un plan normal unique Π que l'on désignera par Π_t.

Si les coordonnées x, y, z sont choisies de telle sorte que le vecteur $\mathbf{t}$ est dirigé suivant l'axe Ox, le plan normal Π_t sera le plan de coordonnées Oxz, et la section normale $\mathscr{L}_t = \Pi \cap \mathscr{X}$ sera définie par l'équation $z = z(x, 0)$ (en coordonnées x, z dans le plan Oxz). Ceci prouve que toute section normale est localement (dans un voisinage de p) un graphique et, par suite, un arc régulier simple.

Toute tangente au graphique $z = z(x, 0)$ est située de toute évidence dans le plan tangent à la surface $z = z(x, y)$ et en même temps dans le plan normal. Donc, elle est dirigée par le vecteur $\mathbf{t}$, i.e. le vecteur tangent à la section normale en p est proportionnel au vecteur $\mathbf{t}$.

Puisque le vecteur $\mathbf{t}$ est défini par la section normale à une constante multiplicative près, on peut, sans perdre en généralité, admettre que c'est un vecteur de base (i.e. de longueur 1). En paramétrant la section normale par un paramètre naturel, on peut par conséquent admettre que le vecteur de base $\mathbf{t}$ est tangent à la section normale en p.

Si l'on traite la section normale comme une courbe du plan Π_t, on peut parler de sa courbure relative k_{rel} en p pour l'orientation du plan Π_t définie par le bivecteur $\mathbf{t} \wedge \mathbf{n}$. En désignant cette courbure



Section normale

par k (t), on définit *ipso facto* une fonction $\mathbf{t} \mapsto k$ (t) sur les vecteurs de base de l'espace tangent $T_p\mathscr{X}$. Trouvons l'expression de cette fonction à l'aide des coordonnées de t.

* * *

Soient

(1) $$u = u(s), \quad v = v(s), \quad |s| < s_0$$

les équations paramétriques d'une section normale $\mathscr{L}_t$ de la surface $\mathscr{X}$, où s est un paramètre naturel mesuré à partir du point p. En tant que courbe gauche, $\mathscr{L}_t$ est définie par une équation paramétrique vectorielle $\mathbf{r} = \mathbf{r}(u(s), v(s))$, donc, son vecteur tangent $\dot{\mathbf{r}}$ est donné par la formule

(2) $$\dot{\mathbf{r}} = \mathbf{r}_u \dot{u} + \mathbf{r}_v \dot{v},$$

où comme toujours $\dot{u}$ et $\dot{v}$ sont les dérivées des fonctions (1) par rapport à s.

La dérivée $\ddot{\mathbf{r}}$ du vecteur (2) lui est orthogonale et est parallèle au plan Π_t pour $s = 0$. Donc, elle est proportionnelle au vecteur $\mathbf{n}$ pour $s = 0$. Le coefficient de proportionnalité correspondant (qui est égal au produit scalaire $\ddot{\mathbf{r}}\mathbf{n}$) n'est autre (cf. définition de la courbure relative à la leçon (2 que la courbure relative k (t). Donc,

(3) $$k(\mathbf{t}) = \ddot{\mathbf{r}}\mathbf{n}.$$

Puisque

$$\ddot{\mathbf{r}} = \dot{\mathbf{r}}_u \dot{u} + \mathbf{r}_u \ddot{u} + \dot{\mathbf{r}}_v \dot{v} + \mathbf{r}_v \ddot{v} =$$

$$= (\mathbf{r}_{uu}\dot{u} + \mathbf{r}_{uv}\dot{v})\,\dot{u} + (\mathbf{r}_{uv}\dot{u} + \mathbf{r}_{vv}\dot{v})\,\dot{v} + \mathbf{r}_u\ddot{u} + \mathbf{r}_v\ddot{v} =$$

$$= \mathbf{r}_{uu}\dot{u}^2 + 2\mathbf{r}_{uv}\dot{u}\dot{v} + \mathbf{r}_{vv}\dot{v}^2 + \mathbf{r}_u\ddot{u} + \mathbf{r}_v\ddot{v}$$

et $\mathbf{r}_u\mathbf{n} = 0$, $\mathbf{r}_v\mathbf{n} = 0$, ceci prouve que

$$k(\mathbf{t}) = L\dot{u}^2 + 2M\dot{u}\dot{v} + N\dot{v}^2,$$

où

(4) $$L = \mathbf{r}_{uu}\mathbf{n}, \quad M = \mathbf{r}_{uv}\mathbf{n}, \quad N = \mathbf{r}_{vv}\mathbf{n}.$$

Il est commode de prolonger la fonction $\mathbf{t} \longmapsto k(\mathbf{t})$ à tous les vecteurs tangents $\mathrm{d}\mathbf{r} \neq 0$ en admettant par définition que

$$k(\mathrm{d}\mathbf{r}) = k\left(\frac{\mathrm{d}\mathbf{r}}{|\mathrm{d}\mathbf{r}|}\right).$$

Comme $|\mathrm{d}\mathbf{r}|^2 = \mathrm{d}\mathbf{r}^2 = \mathrm{d}s^2$, où

$$\mathrm{d}s^2 = E\,\mathrm{d}u^2 + 2F\,\mathrm{d}u\,\mathrm{d}v + G\,\mathrm{d}v^2$$

(cf. formule (28) de la leçon 3) et que les coordonnées $\dot{u}$ et $\dot{v}$ du vecteur $\frac{\mathrm{d}\mathbf{r}}{|\mathrm{d}\mathbf{r}|} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}$ sont égales à $\frac{\mathrm{d}u}{\mathrm{d}s}$ et $\frac{\mathrm{d}v}{\mathrm{d}s}$, il vient

(5) $$k(\mathrm{d}\mathbf{r}) = \frac{L\,\mathrm{d}u^2 + 2M\,\mathrm{d}u\,\mathrm{d}v + N\,\mathrm{d}v^2}{E\,\mathrm{d}u^2 + 2F\,\mathrm{d}u\,\mathrm{d}v + G\,\mathrm{d}v^2}.$$

[Remarquons que les égalités $\dot{u} = \frac{\mathrm{d}u}{\mathrm{d}s}$ et $\dot{v} = \frac{\mathrm{d}v}{\mathrm{d}s}$ sont formellement confondues avec celles — classiques en analyse — des dérivées traitées comme des quotients de différentielles. Mais leur signification est différente ici, car $\mathrm{d}u$ et $\mathrm{d}v$ ne sont pas des différentielles, mais les coordonnées du vecteur tangent $\mathrm{d}\mathbf{r}$, et $\mathrm{d}s$ est sa longueur.]

Définition 1. La forme quadratique

$$L\,\mathrm{d}u^2 + 2M\,\mathrm{d}u\,\mathrm{d}v + N\,\mathrm{d}v^2$$

s'appelle *deuxième forme quadratique* de la surface $\mathscr{X}$. On la désignera par le symbole II.

Dans ces notations, la formule (5) s'écrit sous la forme suivante facile à retenir:

(6) $$k = \frac{\mathrm{II}}{\mathrm{I}}.$$

Pour les coefficients de la forme II, outre les formules (4) on a

(7) $$L = -\mathbf{r}_u\mathbf{n}_u, \quad M = -\mathbf{r}_u\mathbf{n}_v = -\mathbf{r}_v\mathbf{n}_u, \quad N = -\mathbf{r}_v\mathbf{n}_v.$$

En effet, puisque $\mathbf{r}_u\mathbf{n} = 0$, il vient $\mathbf{r}_{uu}\mathbf{n} + \mathbf{r}_u\mathbf{n}_u = 0$ et $\mathbf{r}_{uv}\mathbf{n} + \mathbf{r}_u\mathbf{n}_v = 0$, i.e. $L = -\mathbf{r}_u\mathbf{n}_u$ et $M = -\mathbf{r}_u\mathbf{n}_v$. De façon analogue, comme $\mathbf{r}_v\mathbf{n} = 0$, alors $\mathbf{r}_{vu}\mathbf{n} + \mathbf{r}_v\mathbf{n}_u = 0$ et $\mathbf{r}_{vv}\mathbf{n} + \mathbf{r}_v\mathbf{n}_v = 0$, i.e. $M = -\mathbf{r}_v\mathbf{n}_u$ et $N = -\mathbf{r}_v\mathbf{n}_v$. □

En introduisant le vecteur

$$\mathrm{d}\mathbf{n} = \mathbf{n}_u\,\mathrm{d}u + \mathbf{n}_v\,\mathrm{d}v, \tag{8}$$

on peut, grâce à ces formules, mettre II sous la forme

$$\mathrm{II} = -\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{n}$$

qui est analogue à l'écriture $\mathrm{I} = \mathrm{d}\mathbf{r}^2$ de la forme I.

Remarque 1. Par analogie, on peut introduire aussi la *troisième forme quadratique* $\mathrm{III} = \mathrm{d}\mathbf{n}^2$. Mais nous verrons qu'elle n'apporte rien de nouveau, car on montrera qu'elle s'exprime linéairement en fonction des formes I et II.

Les coefficients L, M, N de la forme II sont désignés aussi par D, D', D''.

* * *

Pour représenter la fonction $\mathbf{t} \mapsto k(\mathbf{t})$ de façon suggestive, le mathématicien français Dupin a proposé de considérer sur le plan tangent la courbe (appelée aujourd'hui *indicatrice de Dupin*) qui est le lieu géométrique de l'extrémité du vecteur parallèle et de même sens que le vecteur tangent unitaire, d'origine le point de tangence p (pris pour origine des coordonnées dans le plan tangent) et de longueur $|k(\mathbf{t})|^{-1/2}$. Désignons par x et y les coordonnées de cette extrémité (dans le système $O\mathbf{r}_u\mathbf{r}_v$); sa longueur s'exprimera alors par la formule

$$|x\mathbf{r}_u + y\mathbf{r}_v| = \sqrt{\mathrm{I}(x, y)}.$$

Puisque la courbure $k(\mathbf{t})$ est donnée par la formule (6) qui dans les nouvelles notations s'écrit

$$k(\mathbf{t}) = \frac{\mathrm{II}(x, y)}{\mathrm{I}(x, y)},$$

on obtient l'équation suivante de l'indicatrice de Dupin

$$\sqrt{\mathrm{I}(x, y)} = \sqrt{\frac{\mathrm{I}(x, y)}{\mathrm{II}(x, y)}},$$

i.e. l'équation

$$|\mathrm{II}(x, y)| = 1.$$

Ceci prouve que *l'indicatrice de Dupin est une courbe d'équation*

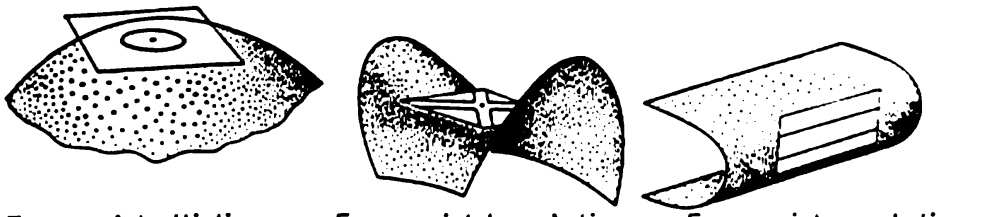
$$|Lx^2 + 2Mxy + Ny^2| = 1.$$

Pour $LN - M^2 > 0$, cette courbe (ou, plus exactement, l'ensemble de ses points réels qui seuls nous intéressent) est l'ellipse d'équation

$$Lx^2 + 2Mxy + Ny^2 = \varepsilon, \tag{9}$$

où $\varepsilon = +1$ si $L > 0$, et $\varepsilon = -1$ si $L < 0$. Pour cette raison, le point de la surface en lequel $LN - M^2 > 0$ est dit *elliptique*.

En un point elliptique, les courbures k (t) sont toutes de même signe (le signe de L). Parmi ces courbures il en existe une plus grande k_1 et une plus petite k_2 (si seulement elles ne sont pas confondues,



En un point elliptique *En un point hyperbolique* *En un point parabolique*

Indicatrice de Dupin

i.e. si seulement l'indicatrice de Dupin n'est pas un cercle), correspondant aux directions des petit et grand axes de l'ellipse (9).

Si $LN - M^2 < 0$, l'indicatrice de Dupin est constituée de deux hyperboles

$$Lx^2 + 2Mxy + Ny^2 = \pm 1 \tag{10}$$

d'asymptotes communes. C'est pourquoi le point de la surface en lequel $LN - M^2 < 0$ est dit *hyperbolique*. La courbure k (t) prend sa valeur maximale $k_1 > 0$ dans la direction de l'axe transverse de l'une des hyperboles (10). Lorsque le vecteur tourne, la courbure commence par diminuer jusqu'à 0 (la direction du vecteur **t** est alors confondue avec celle d'une asymptote), puis elle continue de diminuer jusqu'à sa plus petite valeur $k_2 < 0$ (la direction du vecteur **t** est alors confondue avec celle de l'axe transverse de l'autre hyperbole, i.e. avec celle de l'axe non transverse de la première hyperbole).

Au point en lequel $LN - M^2 = 0$ et qui est dit *parabolique*, l'indicatrice de Dupin est définie par l'équation

$$(\sqrt{|L|}\,x + \sqrt{|N|}\,y)^2 = 1; \tag{11}$$

on reconnaît ici un couple de droites parallèles (pourvu que $L \neq 0$ ou $N \neq 0$). Dans la direction de ces droites, la courbure k (t) est nulle ; dans la direction perpendiculaire, elle atteint sa plus grande valeur (en module) en gardant le même signe. Si $L = 0$, $N = 0$ (donc, $M = 0$), la courbure k (t) est identiquement nulle par rapport à **t** (et l'indicatrice de Dupin n'est pas définie).

Remarquons que l'indicatrice de Dupin est une conique aux points elliptiques et paraboliques, et une courbe d'ordre quatre aux points hyperboliques.

Dans chacun de ces trois cas, la fonction k (t) (pourvu qu'elle ne soit pas identiquement nulle) atteint *deux* fois son maximum k_1 et son minimum k_2.

Exercice 1. Montrer que

$$k(\mathbf{t}) = k_1 \cos^2 \varphi + k_2 \sin^2 \varphi,$$

où φ est l'angle du vecteur t et de la direction dans laquelle la courbure est égale à k_1. Cette formule s'appelle f o r m u l e d'E u l e r.

* * *

Définition 2. Les nombres k_1 et k_2 s'appellent *courbures principales* de la surface au point envisagé. Leur produit

$$K = k_1 k_2$$

s'appelle *courbure totale* (ou *gaussienne*) et leur demi-somme

$$H = \frac{k_1 + k_2}{2}$$

courbure moyenne.

D'après ce qui précède, $K > 0$ en un point elliptique, $K < 0$ en un point hyperbolique, et $K = 0$ en un point parabolique.

Soulignons que H et K dépendent du point $p \in \mathscr{X}$, i.e. sont des fonctions sur $\mathscr{X}$. Ces fonctions sont différentiables par rapport aux coordonnées locales.

Pour trouver les courbures principales, on aurait pu chercher les directions principales des coniques (9) et (10) (pour la courbe (11) le problème ne se pose pas) et former ensuite leurs équations canoniques. Malheureusement, cette voie implique de longs calculs à cause de la non-rectangularité des coordonnées x et y. Nous allons donc procéder d'une autre façon en nous adressant directement à la formule fondamentale (6).

D'après cette formule, la courbure k_2 est un minimum de la fonction

$$\frac{\mathrm{II}(x, y)}{\mathrm{I}(x, y)} = \frac{Lx^2 + 2Mxy + Ny^2}{Ex^2 + 2Fxy + Gy^2}$$

des variables x et y pour $(x, y) \neq (0, 0)$. Donc,

$$\frac{\mathrm{II}(x, y)}{\mathrm{I}(x, y)} \geqslant k_2$$

pour tout $(x, y) \neq (0, 0)$, l'égalité étant réalisée en un point (x, y) au moins. Comme $\mathrm{I}(x, y) > 0$ pour $(x, y) \neq (0, 0)$, cette inégalité

équivaut à l'inégalité suivante

$$\mathrm{II}\,(x, y) - k_2\mathrm{I}\,(x, y) \geqslant 0$$

qui exprime que la forme quadratique $\mathrm{II} - k_2\,\mathrm{I}$ de matrice

$$\begin{Vmatrix} L-k_2E & M-k_2F \\ M-k_2F & N-k_2G \end{Vmatrix}$$

est positive en tous les points $(x, y) \neq (0, 0)$ et nulle au moins en l'un d'eux.

De façon analogue, le nombre k_1 est caractérisé par le fait que la forme quadratique $\mathrm{II} - k_1\mathrm{I}$ est positive en tous les points $(x, y) \neq (0, 0)$ et nulle en l'un d'eux au moins.

Mais il est aisé de voir (immédiatement ou en vertu de la théorie générale des formes quadratiques sur le corps $\mathbb{R}$; cf. leçon II.12) que *la forme quadratique de deux variables est négative ou positive en tous les points* $(x, y) \neq (0, 0)$ *et nulle en l'un d'eux au moins si et seulement si son rang est* < 2, *i.e. le déterminant de sa matrice est nul.*

Ceci prouve que *les courbures principales* k_1 *et* k_2 *sont les racines de l'équation*

$$\begin{vmatrix} L-kE & M-kF \\ M-kF & N-kG \end{vmatrix} = 0,$$

i.e. de l'équation

$$(EG - F^2)k^2 - (EN + GL - 2FM)\,k + (LN - M^2) = 0.$$

D'après les formules de Viète on en déduit en particulier que

$$K = \frac{LN-M^2}{EG-F^2}, \quad H = \frac{1}{2}\,\frac{EN+GL-2FM}{EG-F^2}.$$

La première de ces formules trouvera une importante application à la leçon suivante.

* * *

Supposons que l'espace $\mathcal{A}$ est muni de coordonnées x, y, z dans lesquelles la surface étudiée est le graphique d'une fonction $z = z\,(x, y)$, et de plus $z\,(0, 0) = 0$ et le vecteur normal au point $(0, 0)$ est le vecteur unitaire $\mathbf{k}$ de l'axe Oz (cf. plus haut).

Alors au voisinage du point $(0, 0)$, les coordonnées x et y sont des coordonnées u et v sur la surface et de plus

$$\mathbf{r}_u = (1, 0, z_x), \qquad \mathbf{r}_v = (0, 1, z_y).$$

Puisque les vecteurs $\mathbf{r}_u$ et $\mathbf{r}_v$ sont en $(0, 0)$ parallèles par hypothèse au plan Oxy, il vient

$$(z_x)_0 = 0 \quad \text{et} \quad (z_y)_0 = 0$$

(l'indice 0 désigne les valeurs des dérivées partielles en (0, 0)); donc, le développement de la fonction $z(x, y)$ en série de Taylor commence par les termes du second degré :

$$z = \frac{1}{2}(rx^2 + 2sxy + ty^2) + \ldots,$$

où

$$r = (z_{xx})_0, \quad s = (z_{xy})_0, \quad t = (z_{yy})_0$$

(notations de Monge).

D'autre part, puisque

$$\mathbf{r}_{uu} = (0, 0, z_{xx}), \quad \mathbf{r}_{uv} = (0, 0, z_{xy}), \quad \mathbf{r}_{vv} = (0, 0, z_{yy}),$$

on a au point (0, 0)

$$L = r, \quad M = s, \quad N = t.$$

Donc, dans le cas envisagé, *la deuxième forme quadratique de la surface ne diffère que du facteur constant* $^1/_2$ *de la somme* $z_2(x, y)$ *des termes du second degré de la série de Taylor de* $z(x, y)$.

Puisque dans un voisinage du point (0, 0) la surface $z = z(x, y)$ diffère peu de la surface $z = z_2(x, y)$ et que cette dernière est un paraboloïde elliptique pour $rt - s^2 > 0$ et un paraboloïde hyperbolique pour $rt - s^2 < 0$, ceci prouve qu'*une surface diffère peu d'un paraboloïde elliptique au voisinage d'un point elliptique, et d'un paraboloïde hyperbolique au voisinage d'un point hyperbolique.*

Ceci nous donne une idée assez satisfaisante du comportement d'une surface au voisinage des points non paraboliques.

Lorsque $rt - s^2 = 0$, mais soit $r \neq 0$, soit $s \neq 0$, la surface $z = z_2(x, y)$ est un cylindre parabolique. Donc, *au voisinage d'un point parabolique pour lequel soit* $L \neq 0$, *soit* $N \neq 0$, *la surface diffère peu d'un cylindre parabolique.*

On ne peut rien dire de précis sur la structure de la surface au voisinage d'un point parabolique en lequel $L = 0$ et $N = 0$ (donc, $M = 0$) ; en général, elle peut être très complexe.

Les calculs effectués montrent aussi qu'au point (0, 0) on a les égalités $E = 1$, $F = 0$ et $G = 0$ pour la surface envisagée, d'où il résulte qu'en ce point

$$K = rt - s^2 \quad \text{et} \quad H = \frac{r+t}{2}.$$

D'autre part, un calcul immédiat (que l'on peut faire mentalement si l'on remarque que la dérivée $(fg)'$ du produit fg de fonctions f et g telles que $f(0) = 0$ et $g(0) = 1$ prend en 0 la valeur $f'(0)$) montre qu'au point (0, 0)

$$\mathbf{n}_u = (-r, -s, 0), \quad \mathbf{n}_v = (-s, -t, 0),$$

donc,

$$\mathbf{n}_u^2 = r^2 + s^2, \quad \mathbf{n}_u\mathbf{n}_v = s(r+t), \quad \mathbf{n}_v^2 = s^2 + t^2.$$

Il s'ensuit que

$$\mathbf{n}_u^2 = 2HL - KE, \quad \mathbf{n}_u\mathbf{n}_v = 2HM - KF, \quad \mathbf{n}_v^2 = 2HN - KG,$$

i.e. que

$$\mathrm{III} = 2H\mathrm{II} - K\mathrm{I},$$

où III est la troisième forme quadratique introduite dans la remarque 1. Donc, la forme III est bien une combinaison linéaire des formes I et II.

* * *

Pour la surface réglée

$$\mathbf{r} = \boldsymbol{\rho}(u) + v\mathbf{a}(u), \tag{12}$$

on sait déjà que

$$E = 1 + 2v\dot{\boldsymbol{\rho}}\mathbf{a} + v^2\dot{\mathbf{a}}^2, \quad F = \dot{\boldsymbol{\rho}}\mathbf{a}, \quad G = 1$$

(comme toujours on admet que la courbe $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$ est paramétrée par u naturel et que le vecteur $\mathbf{a}(u)$ est unitaire). Ensuite,

$$\mathbf{r}_u = \dot{\boldsymbol{\rho}} + v\dot{\mathbf{a}}, \quad \mathbf{r}_v = \mathbf{a},$$

$$\mathbf{r}_u \times \mathbf{r}_v = \dot{\boldsymbol{\rho}} \times \mathbf{a} + v(\dot{\mathbf{a}} \times \mathbf{a}),$$

$$\mathbf{n} = \frac{\dot{\boldsymbol{\rho}} \times \mathbf{a} + v(\dot{\mathbf{a}} \times \mathbf{a})}{\sqrt{EG - F^2}},$$

$$\mathbf{r}_{uu} = \ddot{\boldsymbol{\rho}} + v\ddot{\mathbf{a}}, \quad \mathbf{r}_{uv} = \dot{\mathbf{a}}, \quad \mathbf{r}_{vv} = 0,$$

$$L = \frac{(\ddot{\boldsymbol{\rho}} + v\ddot{\mathbf{a}})(\dot{\boldsymbol{\rho}} \times \mathbf{a} + v(\dot{\mathbf{a}} \times \mathbf{a}))}{\sqrt{EG - F^2}}, \quad M = \frac{\dot{\boldsymbol{\rho}}\mathbf{a}\dot{\mathbf{a}}}{\sqrt{EG - F^2}}, \quad N = 0,$$

$$LN - M^2 = -\frac{(\dot{\boldsymbol{\rho}}\mathbf{a}\dot{\mathbf{a}})^2}{EG - F^2},$$

donc,

$$K = -\frac{(\dot{\boldsymbol{\rho}}\mathbf{a}\dot{\mathbf{a}})^2}{(EG - F)^2} \leqslant 0.$$

Par conséquent, la courbure totale d'une surface réglée est négative en tout point de cette dernière, i.e. *une surface réglée ne présente pas de points elliptiques.*

Lorsque la surface est un cylindre ($\dot{\mathbf{a}} = 0$), un cône ($\mathbf{a} = \boldsymbol{\rho}$,

donc $\dot{\mathbf{a}} = \dot{\boldsymbol{\rho}}$) ou la surface des tangentes ($\mathbf{a} = \dot{\boldsymbol{\rho}}$), la formule obtenue entraîne $K = 0$. Donc, *la courbure totale de toute surface développable est égale à* 0 (en l'un quelconque de ses points).

Réciproquement, si $K = 0$, on a $\dot{\boldsymbol{\rho}}\mathbf{a}\dot{\mathbf{a}} = 0$, i.e. les vecteurs $\dot{\boldsymbol{\rho}}$, $\mathbf{a}$, $\dot{\mathbf{a}}$ sont coplanaires. Si le vecteur $\dot{\mathbf{a}}(u)$ n'est pas identiquement nul, i.e. si la surface (12) n'est pas un cylindre, quitte à passer à un plus petit voisinage, on peut admettre que $\dot{\mathbf{a}}(u) \neq 0$ pour tout u. Donc, les vecteurs $\mathbf{a}$ et $\dot{\mathbf{a}}$ sont linéairement indépendants (ils sont orthogonaux et non nuls) et, par suite, $\dot{\boldsymbol{\rho}}$ est leur combinaison linéaire :

$$\dot{\boldsymbol{\rho}} = \lambda\mathbf{a} + \mu\dot{\mathbf{a}},$$

où $\lambda = \lambda(u)$ et $\mu = \mu(u)$ sont des fonctions de u.

Soient

$$u_1 = u, \qquad v_1 = v + \mu(u).$$

Le jacobien de cette transformation étant égal à 1, les nombres u_1 et v_1 sont aussi, quitte à passer à un plus petit voisinage, des coordonnées sur la surface (12). En ces coordonnées, l'équation de la surface est de la forme

$$\mathbf{r} = \boldsymbol{\rho}_1(u) + v\mathbf{a}(u)$$

(nous omettons les indices de u_1 et v_1), où

$$\boldsymbol{\rho}_1(u) = \boldsymbol{\rho}(u) - \mu(u)\,\mathbf{a}(u),$$

donc,

$$\dot{\boldsymbol{\rho}}_1 = \dot{\boldsymbol{\rho}} - \dot{\mu}\mathbf{a} - \mu\dot{\mathbf{a}} = (\lambda - \dot{\mu})\,\mathbf{a}.$$

Si $\dot{\boldsymbol{\rho}}_1 = 0$ identiquement (i.e. $\lambda = \dot{\mu}$), l'équation de la surface devient

$$\mathbf{r} = \text{const} + v\mathbf{a}(u),$$

donc, cette surface est un cône. Dans le cas contraire, on peut admettre, quitte encore à passer à un plus petit voisinage, que $\dot{\boldsymbol{\rho}}_1(u) \neq 0$ pour tout u. En passant à un paramètre naturel (et en changeant au besoin le signe de v), on trouve que $\dot{\boldsymbol{\rho}}_1 = \mathbf{a}$, i.e. que la surface étudiée est une surface des tangentes.

Nous avons ainsi prouvé la

Proposition 1. *Une condition nécessaire et suffisante pour qu'une surface réglée possède en chacun de ses points une courbure totale nulle est qu'elle soit développable.* □

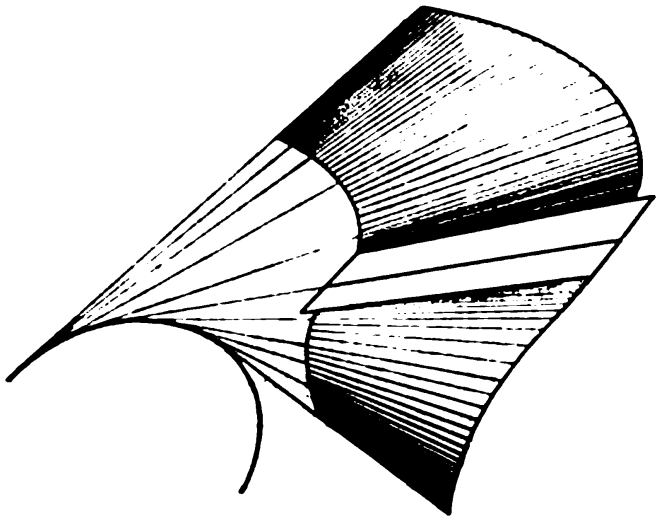
Nous avons établi incidemment que les surfaces développables sont caractérisées par la condition

$$\dot{\boldsymbol{\rho}}\mathbf{a}\dot{\mathbf{a}} = 0,$$

qui équivaut de toute évidence à la colinéarité des vecteurs $\dot{\rho} \times \times \mathbf{a}$ et $\dot{\mathbf{a}} \times \mathbf{a}$. Or, cette colinéarité exprime que le vecteur

$$\mathbf{r}_u \times \mathbf{r}_v = \dot{\rho} \times \mathbf{a} + v\,(\dot{\mathbf{a}} \times \mathbf{a})$$

ne dépend pas de v à un coefficient de proportionnalité près, i.e. le vecteur unitaire correspondant $\mathbf{n}$ ne dépend pas de v. Ceci prouve



Surface développable des tangentes

qu'*une surface développable est une surface réglée qui, en tout point de chacune de ses génératrices rectilignes, admet le même plan tangent.*

* * *

Pour une surface de révolution arbitraire

$$\mathbf{r} = x\,(v) \cos u \cdot \mathbf{i} + x\,(v) \sin u \cdot \mathbf{j} + z\,(v) \cdot \mathbf{k}$$

on a

$$\mathbf{r}_u = -\,x\,(v) \sin u \cdot \mathbf{i} + x\,(v) \cos u \cdot \mathbf{j},$$
$$\mathbf{r}_v = x'\,(v) \cos u \cdot \mathbf{i} + x'\,(v) \sin u \cdot \mathbf{j} + z'\,(v)\,\mathbf{k},$$

donc, $E = x\,(v)^2$, $F = 0$, $G = 1$ (on admet que $x'\,(v)^2 + z'\,(v)^2 = 1$; cf. leçon 3). Par conséquent,

$$\mathbf{r}_u \times \mathbf{r}_v = x\,(v)\,z'\,(v) \cos u \cdot \mathbf{i} + x\,(v)\,z'\,(v) \sin u \cdot \mathbf{j} - x\,(v)\,x'\,(v)\,\mathbf{k},$$
$$\mathbf{n} = z'\,(v) \cos u \cdot \mathbf{i} + z'\,(v) \sin u \cdot \mathbf{j} - x'\,(v)\,\mathbf{k};$$
$$\mathbf{r}_{uu} = -\,x\,(v) \cos u \cdot \mathbf{i} - x\,(v) \sin u \cdot \mathbf{j},$$

$$\mathbf{r}_{uv} = -x'(v)\sin u\cdot\mathbf{i} + x'(v)\cos u\cdot\mathbf{j},$$
$$\mathbf{r}_{vv} = x''(v)\cos u\cdot\mathbf{i} + x''(v)\sin u\cdot\mathbf{j} + z''(v)\,\mathbf{k};$$
$$L = \mathbf{r}_{uu}\mathbf{n} = -x(v)\,z'(v), \quad M = \mathbf{r}_{uv}\mathbf{n} = 0,$$
$$N = \mathbf{r}_{vv}\mathbf{n} = x''(v)\,z'(v) - z''(v)\,x'(v) = -\begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix},$$
$$\frac{LN-M^2}{EG-F^2} = \frac{z'(v)}{x(v)}\begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix}.$$

Ceci prouve que pour une surface de révolution,

$$K = \frac{z'(v)}{x(v)}\begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix}.$$

Exemple 1. Pour une sphère de rayon R on a

$$x(v) = R\cos\frac{v}{R}, \quad z(v) = R\sin\frac{v}{R},$$

donc,

$$x'(v) = -\sin\frac{v}{R}, \quad z'(v) = \cos\frac{v}{R},$$
$$x'' = -\frac{1}{R}\cos\frac{v}{R}, \quad z''(v) = -\frac{1}{R}\sin\frac{v}{R},$$
$$K = \frac{z'(v)}{x(v)}\begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix} = \frac{1}{R^2}.$$

Par conséquent, *la courbure totale d'une sphère de rayon R est constante et égale à $1/R^2$.*

Un résultat qui est géométriquement évident.

L'exemple suivant est plus intéressant.

Exemple 2. La surface de révolution de profil

$$x(v) = R\sin v, \quad z(v) = R\left(\ln\operatorname{tg}\frac{v}{2} + \cos v\right),$$
$$0 < v < \pi/2,$$

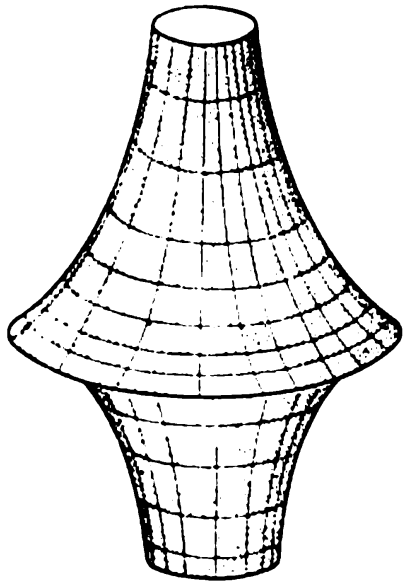
(la *tractrice*) s'appelle *pseudo-sphère* de *pseudo-rayon* R. Pour cette surface,

$$x'(v) = R\cos v, \quad z'(v) = \frac{R}{\sin v} - R\sin v = R\frac{\cos^2 v}{\sin v},$$

donc,

$$x'(v)^2 + z'(v)^2 = R^2\operatorname{cotg}^2 v.$$

Puisque $x'(v)^2 + z'(v)^2 \neq 1$, la formule générale obtenue ci-dessus ne peut être utilisée directement, d'où la nécessité de passer préalablement au paramètre naturel du profil.



Pseudo-sphère

On a

$$s = -R \int_{\pi/2}^{v} \operatorname{cotg} v \, dv = -R \ln \sin v,$$

et, par suite,

$$\sin v = e^{-\frac{s}{R}}, \quad \cos v = \sqrt{1 - e^{-2\frac{s}{R}}},$$

$$\operatorname{tg} \frac{v}{2} = e^{\frac{s}{R}} - \sqrt{e^{\frac{s}{R}} - 1}.$$

Donc, la tractrice est définie par les équations paramétriques

$$x(v) = Re^{-\frac{v}{R}},$$

$$z(v) = R \ln\left(e^{\frac{v}{R}} - \sqrt{e^{2\frac{v}{R}} - 1}\right) + R\sqrt{1 - e^{-2\frac{v}{R}}}.$$

On a

$$x'(v) = -e^{-\frac{v}{R}}, \quad z'(v) = -\sqrt{1 - e^{-2\frac{v}{R}}},$$

$$x''(v) = \frac{1}{R} e^{-\frac{v}{R}}, \quad z''(v) = -\frac{e^{-2\frac{v}{R}}}{R\sqrt{1 - e^{-2\frac{v}{R}}}},$$

$$\begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix} = -\frac{e^{-\frac{v}{R}}}{R\sqrt{1-e^{-2\frac{v}{R}}}},$$

$$\frac{z'(v)}{x(v)}\begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix} = -\frac{1}{R^2}.$$

Donc,

$$K = -\frac{1}{R^2},$$

de sorte que *la courbure totale de la pseudo-sphère est constante et égale à* $-1/R^2$.

Nous voyons que la courbure totale de la pseudo-sphère diffère seulement par son signe de celle de la sphère. C'est à cette propriété que la pseudo-sphère doit son nom.

Exemple 3. Pour la caténoïde,

$$x(v) = \operatorname{ch} v, \quad z(v) = v,$$
$$x'(v) = \operatorname{sh} v, \quad z'(v) = 1,$$
$$x'(v)^2 + z'(v)^2 = \operatorname{ch}^2 v,$$

et il nous faut encore passer au paramètre naturel

$$s = \int_0^v \operatorname{ch} v \, dv = \operatorname{sh} v.$$

En désignant ce paramètre par v, on obtient les équations paramétriques

$$x(v) = \sqrt{1+v^2}, \quad z(v) = \ln(v + \sqrt{1+v^2}).$$

Donc,

$$x'(v) = \frac{v}{\sqrt{1+v^2}}, \quad z'(v) = \frac{1}{\sqrt{1+v^2}},$$
$$x''(v) = \frac{1}{(1+v^2)^{3/2}}, \quad z''(v) = -\frac{v}{(1+v^2)^{3/2}},$$
$$\begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix} = -\frac{1}{1+v^2},$$

et, par suite,

$$K = -\frac{1}{(1+v^2)^2}.$$

Il est intéressant de comparer les courbures de la caténoïde et de l'hélicoïde qui lui est isométrique.

Pour l'hélicoïde on a l'équation (12) avec

$$\rho(u) = u\cdot\mathbf{k} \quad \text{et} \quad \mathbf{a}(u) = \cos u\cdot\mathbf{i} + \sin u\cdot\mathbf{j}.$$

Donc,

$$\dot{\rho} = \mathbf{k}, \quad \dot{\mathbf{a}} = -\sin u\cdot\mathbf{i} + \cos u\cdot\mathbf{j},$$

$$E = 1 + 2v\dot{\rho}\dot{\mathbf{a}} + v^2\dot{\mathbf{a}}^2 = 1 + v^2,$$

$$F = \dot{\rho}\mathbf{a} = 0, \quad G = 1,$$

$$EG - F^2 = 1 + v^2,$$

$$\dot{\rho}\mathbf{a}\dot{\mathbf{a}} = \begin{vmatrix} 0 & 0 & 1 \\ \cos u & \sin u & 1 \\ -\sin u & \cos u & 0 \end{vmatrix} = 1,$$

et, par suite,

$$K = -\frac{1}{(1+v^2)^2}.$$

Le résultat est le même que pour la caténoïde! Ceci signifie que *lorsqu'on déforme une caténoïde en un hélicoïde, les courbures totales sont égales aux points homologues.*

Qu'en est-il de la courbure moyenne?

Pour la caténoïde, $E = 1 + v^2$, $F = 0$, $G = 1$. De plus,

$$L = -x(v)\,z'(v) = -1, \quad M = 0,$$

$$N = -\begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix} = \frac{1}{1+v^2},$$

donc,

$$EN + GL - 2FM = 0,$$

i.e.

$$H = 0.$$

Donc, *la courbure moyenne de la caténoïde est nulle.*

Pour l'hélicoïde,

$$\dot{\rho}\times\mathbf{a} = \sin u\cdot\mathbf{i} - \cos u\cdot\mathbf{j}, \quad \mathbf{a}\times\mathbf{a} = -\mathbf{k},$$

$$\ddot{\rho} = 0, \quad \ddot{\mathbf{a}} = -\cos u\cdot\mathbf{i} - \sin u\cdot\mathbf{j},$$

$$(\ddot{\rho} + v\mathbf{a})(\dot{\rho}\times\mathbf{a} + v(\dot{\mathbf{a}}\times\mathbf{a})) = 0$$

et de plus, comme déjà vu,

$$E = 1 + v^2, \quad F = 0, \quad G = 1,$$

$$EG - F^2 = 1 + v^2, \quad \dot{\rho}\mathbf{a}\dot{\mathbf{a}} = 0.$$

Donc,

$$L=0, \quad M=\frac{1}{\sqrt{1+v^2}}, \quad N=0,$$

et, par suite,

$$EN + LG - 2FM = 0,$$

i.e.

$$H = 0.$$

La courbure moyenne de l'hélicoïde est nulle aussi!

L'exemple de la caténoïde et de l'hélicoïde nous incite à penser que la déformation (l'isométrie) laisse invariantes les courbures totales et moyennes. Pour la courbure totale, ceci est vrai (et nous le montrerons à la prochaine leçon), pour la courbure moyenne, ceci est faux. En effet, la courbure moyenne du plan est nulle, tandis que celle du cylindre circulaire de rayon R qui se déforme en un plan, vaut $1/2R$.

Nous n'avons pas la possibilité d'expliquer ici pourquoi les courbures moyennes de la caténoïde et de l'hélicoïde sont égales.

LEÇON 5

Formules de dérivation de Weingarten. — Coefficients de connexité. — Théorème de Gauss. — Formule explicite de la courbure totale. — Conditions nécessaires et suffisantes d'isométrie. — Surfaces de courbure constante.

La base mobile $\mathbf{r}_u$, $\mathbf{r}_v$, $\mathbf{n}$ d'une surface arbitraire

$$\mathbf{r} = \mathbf{r}(u, v) \tag{1}$$

est justiciable de formules analogues aux formules de Frenet pour les courbes. Ces formules nous donnent la décomposition des dérivées

$$\mathbf{r}_{uu}, \mathbf{r}_{uv}, \mathbf{r}_{vv}, \mathbf{n}_u, \mathbf{n}_v$$

des vecteurs du repère suivant ce repère.

Puisque $\mathbf{n}^2 = 1$, donc $\mathbf{n}\mathbf{n}_u = 0$, $\mathbf{n}\mathbf{n}_v = 0$, les vecteurs $\mathbf{n}_u$ et $\mathbf{n}_v$ se décomposent uniquement suivant les vecteurs $\mathbf{r}_u$ et $\mathbf{r}_v$, de sorte que

$$\mathbf{n}_u = \alpha\mathbf{r}_u + \beta\mathbf{r}_v,$$
$$\mathbf{n}_v = \alpha_1\mathbf{r}_u + \beta_1\mathbf{r}_v.$$

En multipliant la première de ces formules par $\mathbf{r}_u$ et $\mathbf{r}_v$, on obtient deux relations

$$-L = \mathbf{r}_u\mathbf{n}_u = \alpha\mathbf{r}_u^2 + \beta\mathbf{r}_u\mathbf{r}_v = \alpha E + \beta F,$$
$$-M = \mathbf{r}_v\mathbf{n}_u = \alpha\mathbf{r}_u\mathbf{r}_v + \beta\mathbf{r}_v^2 = \alpha F + \beta G,$$

d'où il résulte que

$$\alpha = \frac{FM - GL}{EG - F^2}, \quad \beta = \frac{FL - EM}{EG - F^2}.$$

On calcule de façon analogue les coefficients de la deuxième formule

$$\alpha_1 = \frac{FN - GM}{EG - F^2}, \quad \beta_1 = \frac{FM - EN}{EG - F^2}.$$

D'autre part, puisque par définition

$$\mathbf{r}_{uu}\mathbf{n} = L, \quad \mathbf{r}_{uv}\mathbf{n} = M, \quad \mathbf{r}_{vv}\mathbf{n} = N$$

et que par hypothèse $\mathbf{r}_u\mathbf{n} = 0$, $\mathbf{r}_v\mathbf{n} = 0$, les coefficients de $\mathbf{n}$ dans les décompositions des vecteurs $\mathbf{r}_{uu}$, $\mathbf{r}_{uv}$, $\mathbf{r}_{vv}$ suivant la base $\mathbf{r}_u$, $\mathbf{r}_v$, $\mathbf{n}$ sont respectivement égaux à L, M et N.

On a donc

$$
(2)\qquad
\begin{aligned}
\mathbf{r}_{uu} &= \Gamma^1_{11}\mathbf{r}_u + \Gamma^2_{11}\mathbf{r}_v + L\mathbf{n},\\
\mathbf{r}_{uv} &= \Gamma^1_{12}\mathbf{r}_u + \Gamma^2_{12}\mathbf{r}_v + M\mathbf{n},\\
\mathbf{r}_{vv} &= \Gamma^1_{22}\mathbf{r}_u + \Gamma^2_{22}\mathbf{r}_v + N\mathbf{n},\\
\mathbf{n}_u &= \frac{FM-GL}{EG-F^2}\,\mathbf{r}_u + \frac{FL-EM}{EG-F^2}\,\mathbf{r}_v,\\
\mathbf{n}_v &= \frac{FN-GM}{EG-F^2}\,\mathbf{r}_u + \frac{FM-EN}{EG-F^2}\,\mathbf{r}_v,
\end{aligned}
$$

où Γ^k_{ij}, $i, j, k = 1, 2$, sont des fonctions de u et v. Naguère, ces fonctions étaient désignées par les symboles $\left\{ {i \atop k} {j \atop} \right\}$ et portaient le nom de *symboles de Christoffel*. Aujourd'hui on les appelle en général *coefficients de connexité*.

Les formules (2) s'appellent *formules de dérivation de Weingarten*.

* * *

Pour calculer les coefficients de connexité Γ^k_{ij}, il nous faut déterminer d'abord les six produits des vecteurs $\mathbf{r}_{uu}$, $\mathbf{r}_{uv}$ et $\mathbf{r}_{vv}$ par les vecteurs $\mathbf{r}_u$ et $\mathbf{r}_v$.

Puisque $\mathbf{r}_u^2 = E$, on a $2\mathbf{r}_{uu}\mathbf{r}_u = E_u$ et $2\mathbf{r}_{uv}\mathbf{r}_u = E_v$, i.e.

$$\mathbf{r}_{uu}\mathbf{r}_u = \frac{1}{2}E_u \quad \text{et} \quad \mathbf{r}_{uu}\mathbf{r}_u = \frac{1}{2}E_v.$$

De façon analogue, puisque $\mathbf{r}_v^2 = G$, il vient

$$\mathbf{r}_{uv}\mathbf{r}_v = \frac{1}{2}G_u \quad \text{et} \quad \mathbf{r}_{vv}\mathbf{r}_v = \frac{1}{2}G_v.$$

En outre, comme $\mathbf{r}_u\mathbf{r}_v = F$, on a $\mathbf{r}_{uu}\mathbf{r}_v + \mathbf{r}_u\mathbf{r}_{uv} = F_u$ et $\mathbf{r}_{uv}\mathbf{r}_v + \mathbf{r}_u\mathbf{r}_{vv} = F_v$, d'où il résulte que

$$\mathbf{r}_{uu}\mathbf{r}_v = F_u - \frac{1}{2}E_v \quad \text{et} \quad \mathbf{r}_{vv}\mathbf{r}_u = F_v - \frac{1}{2}G_u.$$

En multipliant maintenant les trois premières formules (2) par $\mathbf{r}_u$ et $\mathbf{r}_v$, on obtient les six relations

$$
(3)\qquad
\begin{cases}
E\Gamma^1_{11} + F\Gamma^2_{11} = \frac{1}{2}E_u,\\
F\Gamma^1_{11} + G\Gamma^2_{11} = F_u - \frac{1}{2}E_v,
\end{cases}
$$

$$
\begin{cases}
E\Gamma^1_{12} + F\Gamma^2_{12} = \frac{1}{2}E_v,\\
F\Gamma^1_{12} + G\Gamma^2_{12} = \frac{1}{2}G_u,
\end{cases}
$$

$$
\begin{cases}
E\Gamma^1_{22} + F\Gamma^2_{22} = F_v - \frac{1}{2}G_u,\\
F\Gamma^1_{22} + G\Gamma^2_{22} = \frac{1}{2}G_v,
\end{cases}
$$

d'où l'on déduit sans peine les coefficients Γ_{ij}^k. (Les équations admettent des solutions uniques, puisque le déterminant $EG - F^2$ de chaque couple d'équations est non nul.)

On voit que *les coefficients de connexité* Γ_{ij}^k *s'expriment en fonction des coefficients de la première forme quadratique et de leurs dérivées.* Donc, *ils sont invariants par les déformations* (*les isométries*) *de la surface.*

Nous n'aurons pas besoin des expressions explicites des coefficients Γ_{ij}^k en fonction des coefficients de la première forme quadratique.

* * *

Les coefficients des formules de dérivation sont liés par trois relations que l'on obtient en calculant les dérivées partielles $\mathbf{r}_{uuv}$, $\mathbf{r}_{uvv}$ et $\mathbf{n}_{uv}$ de deux manières à l'aide de ces formules. L'une de ces relations a été acquise par Gauss, les deux autres par Peterson, Mainardi et Codazzi. Nous allons établir seulement la relation de Gauss en calculant le coefficient de $\mathbf{r}_v$ dans la décomposition de la dérivée partielle $\mathbf{r}_{uuv}$ suivant les vecteurs $\mathbf{r}_u$, $\mathbf{r}_v$ et $\mathbf{n}$.

Dans ce calcul, nous ne nous intéresserons qu'au coefficient de $\mathbf{r}_v$ et à ses termes qui dépendent des coefficients de la deuxième forme quadratique. Tous les autres termes seront remplacés par des pointillés.

On a

$$\begin{aligned}\mathbf{r}_{uuv} = (\mathbf{r}_{uu})_v = (\Gamma_{11}^1\mathbf{r}_u + \Gamma_{11}^2\mathbf{r}_v + L\mathbf{n})_v = \\ = \ldots + \Gamma_{11}^1\mathbf{r}_{uv} + \ldots + \Gamma_{11}^2\mathbf{r}_{vv} + \ldots + L\mathbf{n}_v = \\ = \ldots + \Gamma_{11}^1(\ldots) + \ldots + \Gamma_{11}^2(\ldots) + \ldots \\ \ldots + L\left(\ldots + \frac{FM-EN}{EG-F^2}\mathbf{r}_v\right) = \left(L\frac{FM-EN}{EG-F^2} + \ldots\right)\mathbf{r}_v + \ldots\end{aligned}$$

De façon analogue,

$$\mathbf{r}_{uuv} = (\mathbf{r}_{uv})_u = (\Gamma_{12}^1\mathbf{r}_u + \Gamma_{22}^2\mathbf{r}_v + M\mathbf{n})_u = \left(M\frac{FL-EM}{EG-F^2} + \ldots\right)\mathbf{r}_v + \ldots$$

Donc,

$$L\frac{FM-EN}{EG-F^2} = M\frac{FL-EM}{EG-F^2} + \ldots,$$

où les pointillés désignent les termes dépendant seulement des coefficients de la première forme quadratique. Or

$$M\frac{FL-EM}{EG-F^2} - L\frac{FM-EN}{EG-F^2} = E\frac{LN-M^2}{EG-F^2} = EK.$$

Puisque $E \neq 0$ (la forme I est définie positive), ceci prouve que *la courbure totale K de la surface s'exprime en fonction des coefficients de la première forme quadratique* (*et de leurs dérivées*). Il s'ensuit que la courbure K est invariante par une déformation. Plus exactement, si f est une isométrie d'une surface $\mathcal{X}$ sur une surface $\mathcal{Y}$, alors

$$K_{\mathcal{Y}} \circ f = K_{\mathcal{X}}, \tag{4}$$

où $K_{\mathcal{X}}$ et $K_{\mathcal{Y}}$ sont les courbures totales des surfaces $\mathcal{X}$ et $\mathcal{Y}$ respectivement. (En effet, si f opère par égalisation des coordonnées, les deux membres de (4) ne diffèrent que par les notations des coordonnées.) Ce résultat mérite bien d'être érigé en théorème:

Théorème 1 (théorème de Gauss). *La courbure totale d'une surface est invariante par les déformations* (*isométries*), *i.e. des surfaces isométriques présentent la même courbure aux points homologues.* □

Ce théorème a tellement émerveillé Gauss qu'il l'a appelé *theorema egregium.*

Du théorème 1 il résulte en particulier qu'*aucune portion aussi petite que l'on veut de sphère ne peut être déformée en un plan.* Donc, aucune carte du globe ne peut en donner une représentation exacte.

* * *

La formule explicite de la courbure K en fonction des coefficients E, F et G de la première forme quadratique est

$$K = -\frac{1}{4(EG-F^2)^2}\begin{vmatrix} E & E_u & E_v \\ F & F_u & F_v \\ G & G_u & G_v \end{vmatrix} - \\ -\frac{1}{2\sqrt{EG-F^2}}\left\{\left(\frac{E_v-F_u}{\sqrt{EG-F^2}}\right)_v - \left(\frac{F_v-G_u}{\sqrt{EG-F^2}}\right)_u\right\}. \tag{5}$$

Les deux autres relations qui s'obtiennent par dérivation des formules de dérivation (et qui s'appellent généralement *formules de Peterson-Codazzi*) s'écrivent

$$2(EG-F^2)(L_v-M_u)-(EN+GL-2FM)(E_v-F_u)+ \\ +\begin{vmatrix} E & E_u & L \\ F & F_u & M \\ G & G_u & N \end{vmatrix} = 0,$$

$$2(EG-F^2)(M_v-N_u)-(EN+GL-2FM)(F_v-G_u)+ \\ +\begin{vmatrix} E & E_v & L \\ F & F_v & M \\ G & G_v & N \end{vmatrix} = 0. \tag{6}$$

La démonstration de ces formules n'implique rien d'autre que patience et attention.

Nous allons prouver la formule (5) seule et ce dans le cas où $E = 1$ et $F = 0$, i.e. lorsque la première forme quadratique de la surface est exprimée par la formule

$$\mathrm{I} = du^2 + G\,dv^2. \tag{7}$$

Les équations (3) des coefficients Γ_{ij}^k s'écrivent alors

$$\Gamma_{11}^1 = 0, \qquad \Gamma_{12}^1 = 0, \qquad \Gamma_{22}^1 = -\frac{1}{2}G_u,$$

$$G\Gamma_{11}^2 = 0, \quad G\Gamma_{12}^2 = \frac{1}{2}G_u, \quad G\Gamma_{22}^2 = \frac{1}{2}G_v,$$

d'où il résulte que

$$\Gamma_{11}^1 = 0, \qquad \Gamma_{12}^1 = 0, \qquad \Gamma_{11}^2 = 0,$$

$$\Gamma_{22}^1 = -\frac{1}{2}G_u, \quad \Gamma_{12}^2 = \frac{1}{2}\frac{G_u}{G}, \quad \Gamma_{22}^2 = \frac{1}{2}\frac{G_v}{G}.$$

Donc,

$$\mathbf{r}_{uu} = L\mathbf{n} \quad \text{et} \quad \mathbf{r}_{uv} = \frac{1}{2}\frac{G_u}{G}\mathbf{r}_v + M\mathbf{n}.$$

Puisque dans le cas envisagé

$$\mathbf{n}_u = -L\mathbf{r}_u - \frac{M}{G}\mathbf{r}_v \quad \text{et} \quad \mathbf{n}_v = -M\mathbf{r}_u - \frac{N}{G}\mathbf{r}_v,$$

on en déduit que

$$\mathbf{r}_{uuv} = (L\mathbf{n})_v = L_v\mathbf{n} + L\left(-M\mathbf{r}_u - \frac{N}{G}\mathbf{r}_v\right) = -LM\mathbf{r}_u - \frac{2N}{G}\mathbf{r}_v + L_v\mathbf{n}$$

et

$$\begin{aligned}\mathbf{r}_{uvu} &= \left(\frac{1}{2}\frac{G_u}{G}\mathbf{r}_v + M\mathbf{n}\right)_u = \frac{1}{2}\left(\frac{G_u}{G}\right)_u\mathbf{r}_v + \\ &\quad + \frac{1}{2}\frac{G_u}{G}\left[\frac{1}{2}\frac{G_u}{G}\mathbf{r}_v + M\mathbf{n}\right] + M\mathbf{n} + M\left(-L\mathbf{r}_u - \frac{M}{G}\mathbf{r}_v\right) = \\ &= -LM\mathbf{r}_u + \left[\frac{1}{2}\left(\frac{G_u}{G}\right)_u + \frac{1}{4}\left(\frac{G_u}{G}\right)^2 - \frac{M^2}{G}\right]\mathbf{r}_v + \\ &\qquad + \left(\frac{1}{2}\frac{G_u}{G}M + M_u\right)\mathbf{n},\end{aligned}$$

par conséquent,

$$-\frac{LM}{G} = \frac{1}{2}\left(\frac{G_u}{G}\right)_u + \frac{1}{4}\left(\frac{G_u}{G}\right)^2 - \frac{M^2}{G},$$

$$L_v = \frac{1}{2}\frac{G_u}{G}M + M_u.$$

La deuxième équation ne nous intéresse pas maintenant (c'est la première des équations (6) pour $E = 1$ et $F = 0$), et la première entraîne (car dans le cas envisagé $K = \frac{1}{G}(LN - M^2)$) la formule

$$K = -\frac{1}{2}\left(\frac{G_u}{G}\right)_u - \frac{1}{4}\left(\frac{G_u}{G}\right)^2,$$

i.e., comme le montre un calcul immédiat, la formule

$$K = -\frac{(\sqrt{G})_{uu}}{\sqrt{G}}, \tag{8}$$

ce qui est confondu avec le résultat de la substitution des valeurs $E = 1$ et $F = 0$ dans la formule (5).

Nous avons ainsi prouvé que *la courbure totale d'une surface de première forme quadratique* (7) *est définie par la formule* (8).

Soit par exemple

$$\mathrm{I} = du^2 + \cos^2 u\, dv^2. \tag{9}$$

Alors $\sqrt{G} = \cos u$ et $(\sqrt{G})_{uu} = -\cos u$. Donc, $K = 1$, ce qui est bien compatible avec le résultat de l'exemple 1 de la leçon 4 (puisque la forme (9) est la première forme quadratique de la sphère de rayon $R = 1$; cf. formule (36) de la leçon 3 dans laquelle, il est vrai, il faudra intervertir u et v).

On montre de façon analogue que la surface de première forme quadratique

$$\mathrm{I} = du^2 + \mathrm{ch}^2 u\, dv^2$$

a une courbure $K = -1$ (cf. exemple 2 de la leçon 4).

Remarque 1. Soulignons que tous ces résultats ont lieu pour des surfaces dans un espace euclidien à *trois* dimensions. Pour les surfaces d'un espace de dimension plus élevée, on peut prendre la formule (5) (ou sa forme particulière (8)) pour définition de la courbure K.

Remarque 2. Pour que six fonctions

$$E, F, G, L, M, N \tag{10}$$

définies sur un ensemble convexe ouvert $U \subset \mathbb{R}^2$ soient les coefficients de la première et de la deuxième forme quadratique d'une surface $\mathbf{r} = \mathbf{r}(u, v)$, il faut bien sûr que, outre les conditions

$$E > 0, \quad EG - F^2 > 0 \tag{11}$$

de définition positive, soient satisfaites les relations (5) et (6) $\left(\text{il est entendu que } K = \frac{LM - M^2}{EG - F^2} \text{ est substitué dans la formule (5)}\right)$. Il s'avère que *ces relations sont suffisantes aussi* (pour qu'existe une surface régulière mais généralement non élémentaire ayant les formes I et II). Bien plus, *les fonctions* (10) (vérifiant les relations (5),

(6) et (11)) *définissent une surface à un déplacement près de l'espace.* Ces propositions sont les analogues bidimensionnels de propositions homologues pour les courbes (cf. théorème 1 de la leçon 2) et se démontrent de la même manière (au lieu du théorème d'existence et d'unicité de la solution d'un système d'équations différentielles linéaires, on se sert du théorème homologue pour les systèmes d'équations linéaires aux dérivées partielles). Nous les prouverons — directement pour les variétés de dimension quelconque — dans le prochain ouvrage des *Leçons de géométrie.*

* * *

Le théorème de Gauss affirme que l'égalité des courbures totales est une condition nécessaire d'isométrie de deux surfaces. Dans le même temps, cette condition, qui est loin d'être suffisante, est tellement forte qu'elle permet d'établir sans peine les conditions suffisantes. Nous ne nous appesantirons pas sur cette question et nous étudierons seulement le cas particulier le plus important du théorème correspondant.

Soit

$$\Delta_1 K = \frac{EK_v^2 - 2FK_uK_v + GK_u^2}{EG - F^2}$$

le premier paramètre différentiel de Beltrami de la fonction K. Si les deux fonctions K et $\Delta_1 K$ de u et v sont *fonctionnellement indépendantes*, i.e. si leur jacobien

$$\begin{vmatrix} \dfrac{\partial K}{\partial u} & \dfrac{\partial K}{\partial v} \\ \dfrac{\partial \Delta_1 K}{\partial u} & \dfrac{\partial \Delta_1 K}{\partial v} \end{vmatrix}$$

est partout non nul, on peut les prendre pour nouvelles coordonnées locales sur la surface. Appelons-les *coordonnées gaussiennes.* Du caractère intrinsèque de l'opérateur Δ_1 (formule (34), leçon 3) et de la formule (4) il s'ensuit immédiatement que pour toute isométrie $f : \mathscr{X} \to \mathscr{Y}$ on a l'égalité

$$\Delta_1 K_{\mathscr{Y}} \circ f = \Delta_1 K_{\mathscr{X}}. \tag{12}$$

Ensemble, les formules (4) et (12) expriment que *toute isométrie est une application opérant par égalisation des coordonnées gaussiennes.* On a donc le théorème suivant :

Théorème 2. *Deux surfaces élémentaires sur lesquelles sont définies des coordonnées gaussiennes sont isométriques si et seulement si leurs premières formes quadratiques sont confondues dans ces coordonnées.* □

Donc, pour savoir si deux surfaces sont isométriques ou non, il faut les munir (dans la mesure du possible) de coordonnées gaus-

siennes et calculer leurs premières formes quadratiques dans ces coordonnées. Si ces formes sont confondues, les surfaces sont isométriques, et *vice versa*.

* * *

Le théorème 2 ne donne pas de réponse lorsque K et $\Delta_1 K$ sont fonctionnellement dépendantes, par exemple dans le cas où $\Delta_1 K = 0$ (ce qui a lieu, ainsi qu'il est aisé de le voir, si et seulement si $K = \text{const}$). Du reste, dans ce cas extrême, la condition du théorème 1 s'avère suffisante, i.e. *deux surfaces élémentaires de courbure totale constante sont isométriques si et seulement si elles possèdent la même courbure*. En d'autres termes, *toute surface de courbure totale constante K est localement isométrique à une sphère de rayon $R = \frac{1}{\sqrt{K}}$ si $K > 0$, à un plan si $K = 0$, et à une pseudo-sphère de pseudo-rayon $R = \frac{1}{\sqrt{-K}}$ si $K < 0$.*

Pour le prouver nous aurons besoin d'un lemme qui sera établi dans le prochain ouvrage de cette série.

Lemme de Gauss. *Sur toute surface il existe des coordonnées locales u, v dans lesquelles sa première forme quadratique est de la forme (7), la fonction $G = G(u, v)$ étant telle que*

$$G(0, v) = 1 \quad \text{et} \quad G_u(0, v) = 0 \text{ pour tout } v. \tag{13}$$

En vertu de ce lemme, on peut, sans perdre en généralité, admettre que la première forme quadratique de la surface envisagée de courbure totale constante K est de la forme (7) et donc que K est justiciable de la formule (8). Cette formule peut être traitée comme une équation différentielle du second ordre à coefficients constants par rapport à la fonction $\sqrt{G}$. La théorie des équations différentielles nous dit que la solution générale de cette équation est de la forme

$$\sqrt{G} = \begin{cases} A \cos a(u+B) & \text{si } K = a^2 > 0, \\ Au + B & \text{si } K = 0, \\ A \operatorname{ch} a(u+B) & \text{si } K = -a^2 < 0, \end{cases}$$

où A et B sont des fonctions arbitraires de v. Mais, compte tenu de la première condition (13), on doit avoir

$$\begin{aligned} A \cos aB &= 1 \quad \text{pour } K > 0, \\ B &= 1 \quad \text{pour } K = 0, \\ A \operatorname{ch} aB &= 1 \quad \text{pour } K < 0; \end{aligned}$$

compte tenu de la deuxième condition (13), les identités

$$(\sqrt{G})_u = \frac{1}{2} \frac{G_u}{\sqrt{G}} = \begin{cases} -Aa \sin a(u+B) & \text{pour } K > 0, \\ A & \text{pour } K = 0, \\ Aa \operatorname{sh} a(u+B) & \text{pour } K < 0 \end{cases}$$

aidant, on doit avoir

$$\begin{aligned} Aa \sin aB &= 0 \quad \text{pour } K > 0, \\ A &= 0 \quad \text{pour } K = 0, \\ Aa \operatorname{sh} aB &= 0 \quad \text{pour } K < 0 \end{aligned}$$

Donc,

$$\sqrt{G} = \begin{cases} \cos au & \text{pour } K = a^2 > 0, \\ 1 & \text{pour } K = 0, \\ \operatorname{ch} au & \text{pour } K = -a^2 < 0. \end{cases}$$

Dans le premier cas, on obtient la première forme quadratique

$$du^2 + \cos^2 au \, dv^2$$

de la sphère de rayon $1/a$, dans le deuxième, la première forme quadratique

$$du^2 + dv^2$$

du plan, et dans le troisième, la première forme quadratique

$$du^2 + \operatorname{ch}^2 au \, dv^2$$

de la pseudo-sphère de pseudo-rayon $1/a$. □

Nous arrêtons ici l'exposé de la théorie des surfaces pour entrer dans le vif du sujet de ce cours : la théorie des variétés différentiables.

LEÇON 6

Remarques préliminaires. — Sous-ensembles ouverts de l'espace $\mathbb{R}^n$ et leurs difféomorphismes. — Cartes et atlas. — Atlas maximaux. — Variétés différentiables. — Exemples de variétés différentiables.

La notion de variété différentiable est fondamentale en mathématiques modernes. Elle est apparue dans le cadre de la définition rigoureuse et de la généralisation de la notion intuitive de surface envisagée indépendamment de sa position dans l'espace. Les principes fondamentaux de cette définition sont empruntés à la cartographie.

Les diverses régions du globe terrestre peuvent être décrites de façon adéquate par des cartes qui les représentent sur un plan. Chaque point du globe peut être représenté sur une carte, mais une carte ne peut à elle seule recouvrir tout le globe ; pour cela il faut un atlas, i.e. un ensemble de plusieurs cartes. Toute carte permet de transporter les coordonnées rectangulaires du plan dans un domaine d'une surface et munir ainsi cette surface de c o o r d o n n é e s l o c a l e s. (En réalité, en cartographie, on fait généralement l'opération inverse, c'est-à-dire qu'on transporte les coordonnées géographiques du globe terrestre en un réseau de coordonnées curvilignes sur le plan, mais cette différence n'est pas fondamentale.) Ceci étant, les coordonnées locales associées à deux cartes différentes sont reliées par des fonctions de passage qui permettent d'exprimer (dans la partie commune aux deux cartes) les unes de ces coordonnées par l'intermédiaire des autres.

Dans les définitions générales respectives, on remplacera le plan par l'espace euclidien standard $\mathbb{R}^n$, où n est un entier que nous choisirons et fixerons une fois pour toutes. Le cas dégénéré $n = 0$ ne sera pas pour l'instant écarté.

* * *

Rappelons tout d'abord quelques faits et définitions d'analyse relatifs à l'espace $\mathbb{R}^n$ (qui ont été partiellement utilisés dans les leçons précédentes).

On dit qu'un point x est un *point intérieur* d'un sous-ensemble $U \subset \mathbb{R}^n$ s'il existe un $\varepsilon > 0$ tel que la boule de rayon ε centrée en x soit entièrement contenue dans U. Un ensemble U est *ouvert* (dans $\mathbb{R}^n$) si tous ses points sont intérieurs.

Une application $\varphi: U \to V$ de sous-ensembles de l'espace $\mathbb{R}^n$ est définie par n fonctions

$$(1) \qquad y^1 = \varphi^1(x^1, \ldots, x^n), \ \ldots, \ y^n = \varphi^n(x^1, \ldots, x^n)$$

de n variables, qui expriment les coordonnées $y^1, \ldots, y^n$ du point $\mathbf{y} = \varphi(\mathbf{x})$ par l'intermédiaire des coordonnées $x^1, \ldots, x^n$ d'un point $\mathbf{x} \in U$. Si l'ensemble U est ouvert, on peut parler des dérivées d'ordre quelconque des fonctions (1) par rapport à $x^1, \ldots, x^n$ en tout point de U.

Une application φ d'un ensemble ouvert U est une *application de classe* C^r, où r est un entier naturel ou le symbole ∞, si en tout point de U les fonctions (1) possèdent pour $r \neq \infty$ des dérivées partielles continues de tous les ordres $\leqslant r$, et pour $r = \infty$, des dérivées partielles de tout ordre. Si les fonctions (1) sont $\mathbb{R}$-analytiques en tout point de U (i.e. se développent en séries entières de rayon de convergence non nul), l'application φ s'appelle *application de classe* C^ω.

Dans la suite, nous fixerons une classe de différentiabilité C^r, $r \leqslant \infty$ ou $r = \omega$, et les applications de cette classe seront appelées tout simplement *applications différentiables*. En principe, tout $r \geqslant 2$ et souvent même $r = 1$ conviendront, mais pour ne pas avoir à veiller à l'apparition de dérivées d'ordre trop élevé, on admettra tacitement, si le contraire n'est pas expressément mentionné, que $r = \infty$ ou $r = \omega$. Du reste, nous discuterons parfois le cas particulier $r = 0$: les applications différentiables seront alors de simples applications continues.

Le cas fondamental sera le cas $r = \infty$, et toutes les situations différentes de $r = \omega$ seront spécialement mentionnées.

Toute application différentiable $\varphi: U \to V$ définit sur U une fonction différentiable

$$D\varphi = \det \left\| \frac{\partial \varphi^i}{\partial x^j} \right\|, \quad i, j = 1, \ldots, n,$$

appelée son *jacobien* (qui est désignée aussi par le symbole J_φ). On sait du cours d'analyse que pour toutes applications différentiables $\varphi: U \to V$ et $\psi: V \to W$, où U, V et W sont des sous-ensembles ouverts de l'espace $\mathbb{R}^n$, l'application composée $\psi \circ \varphi: U \to W$ (qui agit suivant la formule $(\psi \circ \varphi)(\mathbf{x}) = \psi(\varphi(\mathbf{x}))$, $\mathbf{x} \in U$) est différentiable aussi et

$$(2) \qquad D(\psi \circ \varphi)(\mathbf{x}) = (D\psi)(\mathbf{y}) \cdot (D\varphi)(\mathbf{x}), \quad \mathbf{y} = \varphi(\mathbf{x}),$$

pour tout $\mathbf{x} \in U$.

Une application $\varphi: U \to V$ d'ensembles ouverts de $\mathbb{R}^n$ est un *difféomorphisme* si elle est différentiable, bijective et si sa réciproque $\varphi^{-1}: V \to U$ est différentiable aussi. (Pour $r = 0$, toute application bijective continue dont la réciproque est aussi continue s'appelle *homéomorphisme*.)

De la formule (2) il résulte immédiatement que *le jacobien $D\varphi$ de tout difféomorphisme $\varphi: U \to V$ est différent de* 0 *en tout point de U* (et de plus, $(D\varphi^{-1})(\mathbf{y}) = (D\varphi)(\mathbf{x})^{-1}$ pour tout point $\mathbf{y} = \varphi(\mathbf{x})$ de V).

Le théorème de l'application réciproque affirme que *si le jacobien $D\varphi$ d'une application différentiable $\varphi: U \to V$ est différent de* 0 *en un point $\mathbf{x}_0 \in U$, il existe un ensemble ouvert $U' \subset U$ contenant le point $\mathbf{x}_0$ tel que la restriction $\varphi|_{U'}$ de φ à U' est un difféomorphisme de U' sur un ensemble ouvert $V' \subset V$ contenant le point $\mathbf{y}_0 = \varphi(\mathbf{x}_0)$.*

On en déduit en particulier qu'*une application bijective différentiable $\varphi: U \to V$ dont le jacobien est partout non nul est un difféomorphisme.*

Remarque 1. Il existe des applications non bijectives différentiables $\varphi: U \to V$ dont le jacobien est partout non nul. Exemple: l'application φ de la couronne plane $1 < x^2 + y^2 < 2$ sur elle-même, définie par la formule $\varphi(x, y) = (x^2 - y^2, 2xy)$.

* * *

Nous sommes maintenant en mesure de donner les définitions fondamentales de la « cartographie abstraite ».

Soit $\mathscr{X}$ un ensemble arbitraire.

Définition 1. On appelle *carte sur* $\mathscr{X}$ un couple (U, h), où U est un sous-ensemble de $\mathscr{X}$ et h une application bijective de U sur un ensemble ouvert de $\mathbb{R}^n$. L'ensemble U s'appelle *domaine de définition* (ou *support*) de la carte (U, h) et l'application h *application de carte*. Pour tout point $p \in U$, le point $h(p) \in \mathbb{R}^n$ est de la forme $(x^1(p), \ldots, x^n(p))$, où $x^1(p), \ldots, x^n(p) \in \mathbb{R}$. On obtient n fonctions numériques

$$x^1 : p \mapsto x^1(p), \ldots, x^n : p \mapsto x^n(p), p \in U, \tag{3}$$

sur U, appelées *coordonnées locales* de la carte (U, h).

Puisque les coordonnées (3) définissent l'application h de façon unique, au lieu de (U, h) on écrit souvent $(U, x^1, \ldots, x^n)$ et l'application h est appelée *application de coordonnées*.

Remarque sur les notations.

Le symbole $\varphi(x^1, \ldots, x^n)$, où φ est une fonction sur $h(U)$, désignera la fonction

$$\varphi \circ h : p \mapsto \varphi(x^1(p), \ldots, x^n(p)) \text{ sur } U,$$

ce qui est compatible avec la notation traditionnelle d'une fonction composée. Donc la formule

$$y = \varphi(x^1, \ldots, x^n)$$

exprime que y n'est qu'une autre notation de la fonction $\varphi(x^1, \ldots, x^n)$. Pour réduire le nombre de lettres utilisées, on désignera parfois la fonction φ par y et on écrira donc

$$y = y(x^1, \ldots, x^n).$$

Signalons que dans cette formule la lettre y du premier et du second membre désigne des fonctions différentes. Au premier, c'est une fonction sur U, au second, une fonction sur $h(U)$ servant à exprimer une fonction sur U au moyen de $x^1, \ldots, x^n$. Quand on aura à faire une distinction entre ces fonctions, on désignera la première d'entre elles par $y \circ h$.

Soulignons que nous n'utiliserons jamais, comme on le fait généralement, les symboles $\varphi(x^1, \ldots, x^n)$ (ou $y(x^1, \ldots, x^n)$) pour désigner une fonction φ (ou y) sur $h(U)$ (on se le permettra de rares fois pour désigner les valeurs de cette fonction en un point $(x^1, \ldots, x^n) \in h(U)$).

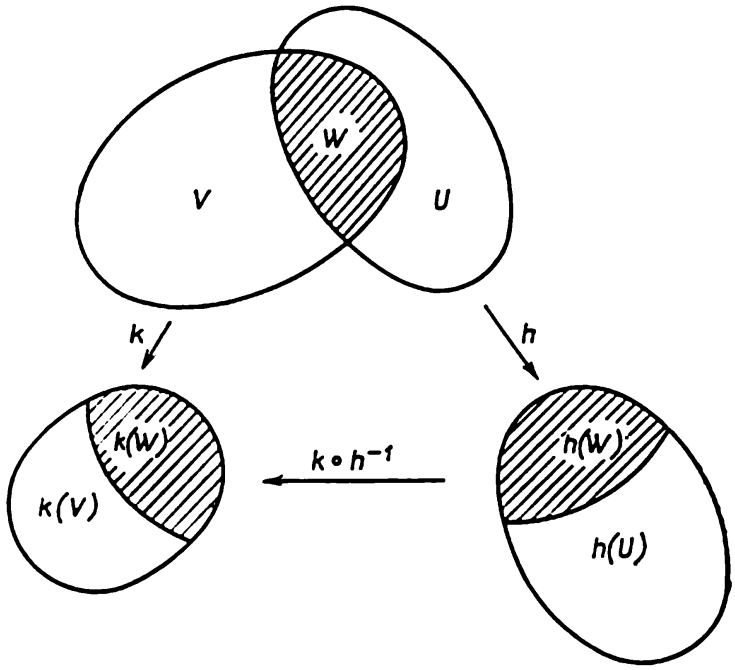
Deux cartes (U, h) et (V, k) sont *sécantes* si $U \cap V \neq \varnothing$. (Cette terminologie reflète la tendance générale, que nous suivrons parfois, à ne pas distinguer par pédantisme (U, h) de U.)

Soient (U, h) et (V, k) deux cartes sécantes et soit $W = U \cap V$. Dans $\mathbb{R}^n$ sont alors définis deux ensembles $h(W)$ et $k(W)$ et une application

$$(k \mid_W) \circ (h \mid_W)^{-1} : h(W) \to k(W) \tag{4}$$

du premier ensemble sur le second. Par abus de notation, on désignera l'application (4) par le symbole $k \circ h^{-1}$.

Définition 2. Deux cartes (U, h) et (V, k) sur $\mathscr{X}$ sont dites *compatibles* si soit elles sont disjointes ($U \cap V = \varnothing$), soit



Pour la suggestion, les ensembles $h(U)$ et $k(V)$ sont pris disjoints, bien qu'ils puissent ne pas l'être dans le cas général. De la même façon, les ensembles $h(U)$, $h(W)$ et $k(V)$ peuvent aussi être non connexes

a) les ensembles $h(W)$ et $k(W)$, où $W = U \cap V$, sont ouverts dans $\mathbb{R}^n$;

b) l'application (4) est un difféomorphisme (pour $r = 0$ un homéomorphisme).

Si $\varphi^1, \ldots, \varphi^n$ sont des fonctions définissant un difféomorphisme (4), les restrictions $x^1 \mid_W, \ldots, x^n \mid_W$ et $y^1 \mid_W, \ldots, y^n \mid_W$ des coordonnées locales des cartes (U, h) et (V, k) à W seront justiciables des formules

$$\begin{aligned} y^1|_W &= \varphi^1(x^1|_W, \ \ldots, \ x^n|_W), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y^n|_W &= \varphi^n(x^1|_W, \ \ldots, \ x^n|_W). \end{aligned} \tag{5}$$

Pour cette raison, les fonctions $\varphi^1, \ldots, \varphi^n$ s'appellent *changements de cartes* ou *fonctions de passage* (sur W) des coordonnées $x^1, \ldots, x^n$

aux coordonnées $y^1, \ldots, y^n$, et les formules (5), *formules de changement de cartes* ou *formules de passage*. Le difféomorphisme $k \circ h^{-1}$ s'appelle *application changement de cartes* ou *application de passage*. D'une façon générale, on omettra l'indice W dans les formules (5) et on écrira

$$(6) \qquad y^1 = \varphi^1(x^1, \ldots, x^n), \ldots, y^n = \varphi^n(x^1, \ldots, x^n)$$

(en ajoutant parfois « sur W »).

On les écrira parfois sous la forme (« vectorielle ») plus concise

$$(7) \qquad \mathbf{y} = \varphi(\mathbf{x}) \quad (\text{ou } \mathbf{y} = (k \circ h^{-1})\,\mathbf{x}).$$

Il faut bien comprendre le caractère conventionnel des formules (7) et ne pas perdre de vue qu'elles ne sont qu'une forme abrégée d'écriture des formules (6) ou (5).

En particulier, les formules (7) ne relient pas des points de $\mathbb{R}^n$, bien qu'elles aient la forme de relations entre les points de sous-ensembles de $\mathbb{R}^n$, mais des fonctions définies sur des sous-ensembles de $\mathscr{X}$ (et en ce sens sont des relations sur $\mathscr{X}$).

Définition 3. On dit qu'un ensemble de cartes $\{(U_\alpha, h_\alpha)\}$ est un *atlas sur* $\mathscr{X}$ si

a) les cartes de cet ensemble sont deux à deux compatibles;

b) on a l'égalité

$$\bigcup_\alpha U_\alpha = \mathscr{X}$$

(les cartes (U_α, h_α) sont un recouvrement de $\mathscr{X}$).

* * *

Pour tout atlas A désignons par $\mathsf{A}_{\max}$ l'ensemble de toutes les cartes compatibles avec chaque carte de l'atlas A.

Proposition 1. *L'ensemble* $\mathsf{A}_{\max}$ *est un atlas.*

Démonstration. Soient (U, h) et (V, k) deux cartes sécantes de $\mathsf{A}_{\max}$ et soit $\mathbf{x}$ un point arbitraire de l'ensemble $h(W)$, où comme plus haut $W = U \cap V$.

Considérons le point $p = h^{-1}(\mathbf{x}) \in W$. L'ensemble $\mathsf{A} = \{(U_\alpha, h_\alpha)\}$ étant un atlas, il existe une carte (U_α, h_α) de A telle que $p \in U_\alpha$. Les cartes (U, h) et (V, k) étant par hypothèse compatibles avec la carte (U_α, h_α), les ensembles $h_\alpha(U_\alpha \cap U)$ et $h_\alpha(U_\alpha \cap V)$ sont ouverts dans $\mathbb{R}^n$. Donc, leur intersection

$$h_\alpha(U_\alpha \cap U) \cap h_\alpha(U_\alpha \cap V) = h_\alpha(U_\alpha \cap W)$$

est ouverte dans $\mathbb{R}^n$ (et par suite dans $h_\alpha(U_\alpha \cap U)$). Par ailleurs, l'ensemble $h(U_\alpha \cap U)$ est ouvert aussi dans $\mathbb{R}^n$ et l'application

$$h \circ h_\alpha^{-1} : h_\alpha(U_\alpha \cap U) \to h(U_\alpha \cap U)$$

est un difféomorphisme (et donc, étant en particulier un homéomorphisme, elle envoie ensemble ouvert dans ensemble ouvert).

Par conséquent, l'ensemble

$$(h \circ h_\alpha^{-1})(h_\alpha(U_\alpha \cap W)) = h(U_\alpha \cap W)$$

est ouvert (dans $h(U_\alpha \cap U)$, donc dans $\mathbb{R}^n$). Puisque $\mathbf{x} \in h(U_\alpha \cap \cap W) \subset h(W)$, ceci prouve que $\mathbf{x}$ est un point intérieur de l'ensemble $h(W)$ et, comme $\mathbf{x}$ est un point arbitraire de $h(W)$, que l'ensemble $h(W)$ est ouvert.

En changeant les rôles de h et k, on établit de façon analogue que $k(W)$ est ouvert aussi.

Ensuite, puisque

$$k \circ h^{-1} = (k \circ h_\alpha^{-1}) \circ (h_\alpha \circ h^{-1}) \tag{8}$$

et que les applications $k \circ h_\alpha^{-1}$ et $h_\alpha \circ h$ sont toutes deux des difféomorphismes, il en est de même de l'application $k \circ h^{-1}$ (en tant que composée de difféomorphismes). Donc, les cartes (U, h) et (V, k) sont compatibles.

[En fait, nous avons conclu assez hâtivement que l'application $k \circ h^{-1}$ est un difféomorphisme, car en toute rigueur la formule (8) n'a pas de sens. En effet, l'application $h_\alpha \circ h^{-1}$ (ou plus exactement l'application $(h_\alpha|_{U_\alpha \cap U}) \circ (h|_{U_\alpha \cap U})^{-1}$) est une application de $(U_\alpha \cap U)$ sur $h_\alpha(U_\alpha \cap U)$, et l'application $k \circ h_\alpha^{-1}$ (plus exactement l'application $(k|_{U_\alpha \cap U}) \circ (h_\alpha|_{U_\alpha \cap U})^{-1}$) une application de $h_\alpha(U_\alpha \cap V)$ sur $k(U_\alpha \cap V)$, donc nous n'avons pas le droit d'envisager la composée de ces applications. Pour acquérir ce droit, il faut restreindre toutes les applications h, k et h_α à $U_\alpha \cap W$, i.e. écrire la formule (8) sous la forme rigoureuse suivante

$$(k|_{U_\alpha \cap W}) \circ (h|_{U_\alpha \cap W})^{-1} =$$
$$= [(k|_{U_\alpha \cap W}) \circ (h_\alpha|_{U_\alpha \cap W})^{-1}] \circ [(h_\alpha|_{U_\alpha \cap W}) \circ (h|_{U_\alpha \cap W})^{-1}].$$

Mais alors on peut seulement affirmer que l'application

$$(k|_{U_\alpha \cap W}) \circ (h|_{U_\alpha \cap W})^{-1} = [(k|_W) \circ (h|_W)^{-1}]|_{h(U_\alpha \cap W)} \tag{9}$$

sera un difféomorphisme, le passage à l'application globale

$$k \circ h^{-1} = (k|_W) \circ (h|_W)^{-1} \tag{10}$$

impliquant des raisonnements complémentaires. Mais ces raisonnements sont assez triviaux. Plus exactement, l'application (9) étant un difféomorphisme, son jacobien est non nul en $\mathbf{x}$ (on admet ici que $r > 0$). Or, il est clair que les jacobiens des applications (9) et (10) prennent la même valeur en $\mathbf{x}$ (car elles sont confondues au voisinage

$h(U_\alpha \cap W)$ de x). Donc, le jacobien de (10) est non nul en tout point x de l'ensemble ouvert $h(W)$. L'application (10) est donc un difféomorphisme, puisqu'elle est différentiable et bijective. (Pour $r = 0$, au lieu de se référer au jacobien, il faut se servir du fait que l'application est continue si et seulement si elle l'est en chaque point.)

Dans la suite, ce genre d'éclaircissements sera laissé au soin du lecteur.]

Nous avons ainsi prouvé la compatibilité de deux cartes quelconques de l'ensemble A_{max}, i.e. que cet ensemble satisfait à la condition a) de la définition 3. Puisque la condition b) de cette définition est visiblement remplie pour A_{max} (et même pour $\mathsf{A} \subset \mathsf{A}_{max}$), la proposition 1 est entièrement prouvée. □

Il est clair que si A et A^* sont des atlas tels que $\mathsf{A} \subset \mathsf{A}^*$, alors $\mathsf{A}^*_{max} \subset \mathsf{A}_{max}$. Donc, en particulier, $(\mathsf{A}_{max})_{max} \subset \mathsf{A}_{max}$ et, par suite, $(\mathsf{A}_{max})_{max} = \mathsf{A}_{max}$. Donc, si $\mathsf{A}_{max} \subset \mathsf{A}^*$, alors $\mathsf{A}^*_{max} \subset (\mathsf{A}_{max})_{max} = \mathsf{A}_{max}$, et $\mathsf{A}^* \subset \mathsf{A}_{max}$. Ce qui exprime que *l'atlas* A_{max} *est maximal* (dans l'ensemble des atlas ordonné par l'inclusion). D'autre part, si $\bar{\mathsf{A}}$ est un atlas maximal arbitraire contenant l'atlas A, alors $\bar{\mathsf{A}} \subset \bar{\mathsf{A}}_{max} \subset \mathsf{A}_{max}$, et, par suite, $\bar{\mathsf{A}} = \mathsf{A}_{max}$. Ceci prouve le

Corollaire 1. *Chaque atlas* A *est contenu dans un seul atlas maximal* A_{max}. □

* * *

Nous sommes maintenant en mesure d'exhiber la définition fondamentale.

Définition 4. Les atlas maximaux sur $\mathscr{X}$ sont aussi appelés *structures différentiables*. Un ensemble $\mathscr{X}$ muni d'une structure différentiable A_{max} s'appelle *variété différentiable*. (Donc, les variétés différentiables sont à proprement parler des couples $(\mathscr{X}, \mathsf{A}_{max})$, mais pour simplifier les notations et les énoncés on ne se servira intégralement de cette notation que par la force des choses.) Les cartes de l'atlas A_{max} s'appellent *cartes de la variété* $\mathscr{X}$ ou même, pour plus d'expressivité, *cartes différentiables*.

Par définition, deux variétés $(\mathscr{X}, \mathsf{A}_{max})$ et $(\mathscr{Y}, \mathsf{A}^*_{max})$ sont *identiques* si et seulement si $\mathscr{X} = \mathscr{Y}$ et $\mathsf{A}_{max} = \mathsf{A}^*_{max}$.

On dit que deux atlas sont *équivalents* s'ils sont contenus dans le même atlas maximal. Il est clair que des atlas A et A^* sont équivalents si et seulement si leur réunion $\mathsf{A} \cup \mathsf{A}^*$ est un atlas (i.e. toute carte de l'un est compatible avec toute carte de l'autre).

Pour définir une structure différentiable A_{max} sur $\mathscr{X}$, il suffit de toute évidence de se donner un atlas arbitraire $\mathsf{A} \subset \mathsf{A}_{max}$. Donc, on peut considérer que les couples $(\mathscr{X}, \mathsf{A})$, où A est un atlas quelconque sur $\mathscr{X}$, sont des variétés différentiables. Ceci étant, des variétés $(\mathscr{X}, \mathsf{A})$ et $(\mathscr{Y}, \mathsf{A}^*)$ seront identiques si et seulement si $\mathscr{X} = \mathscr{Y}$

et les atlas A et A* sont équivalents (i.e. si leur réunion A $\cup$ A* est un atlas).

Signalons que dans la définition d'une variété différentiable intervient le nombre n, dimension de l'espace $\mathbb{R}^n$ contenant les images $h(U)$ des supports des cartes.

Définition 5. Ce nombre s'appelle *dimension* de la variété différentiable $\mathscr{X}$ et se note dim $\mathscr{X}$.

Dans la définition de la variété différentiable $\mathscr{X}$ intervient aussi le nombre $r \geqslant 0$ (ou le symbole ∞ ou ω) qui désigne l'ordre de la classe de différentiabilité des changements de cartes $k \circ h^{-1}$. Cette classe s'appelle *classe de différentiabilité de la variété* $\mathscr{X}$ (on dit aussi que $\mathscr{X}$ est une *variété de classe* C^r). Les variétés de classe C^ω s'appellent aussi *variétés* $\mathbb{R}$*-analytiques.*

Pour $r = 0$, le terme « variété différentiable » est remplacé par *variété topologique.*

Il est évident que toute variété de C^r est automatiquement de classe $C^{r'}$, pour tout $r' < r$ (et en particulier est une variété topologique).

D'après ce qui a été convenu ci-dessus, on admettra généralement que toutes les variétés envisagées sont de classe C^∞.

* * *

Exemples de variétés différentiables.

Exemple 1. Sur l'espace $\mathbb{R}^n$ le couple ($\mathbb{R}^n$, id), où id : $\mathbb{R}^n \to \mathbb{R}^n$ est l'application identique, est une carte, et l'ensemble composé de cette unique carte est un atlas. La structure différentiable correspondante sur $\mathbb{R}^n$ (de classe C^ω) s'appelle *structure standard.* Dans la suite, nous traiterons toujours $\mathbb{R}^n$ comme une variété différentiable munie de la structure différentiable standard. Remarquons que dim $\mathbb{R}^n = n$ pour cette structure.

Exercice 1. Montrer que les couples (U, h), où U est un ensemble ouvert de $\mathbb{R}^n$ et h un difféomorphisme de l'ensemble U sur un ouvert $h(U)$ de $\mathbb{R}^n$, sont des cartes pour la structure différentiable standard sur $\mathbb{R}^n$.

Exemple 2. De façon générale, pour tout ensemble ouvert O de l'espace $\mathbb{R}^n$ le couple (O, id) est une carte, et l'ensemble de cartes composé de cette seule carte est un atlas. On dit de la structure différentiable définie sur O par cet atlas qu'elle est induite par la structure différentiable standard sur $\mathbb{R}^n$. Il est clair que seront des cartes pour la structure induite les cartes (U, h) pour la structure standard, telles que $U \subset O$. Dans la suite on admettra que tous les ensembles ouverts O de $\mathbb{R}^n$ sont des variétés différentiables munies de cette structure différentiable.

Exemple 3. Il existe des structures différentiables non standards

sur l'espace $\mathbb{R}$. Considérons par exemple sur $\mathbb{R}^1 = \mathbb{R}$ une carte $(\mathbb{R}, h_0)$, où $h_0 : \mathbb{R} \to \mathbb{R}$ est une application définie par $h_0(t) = t^3$, $t \in \mathbb{R}$. Cette carte constitue aussi, à elle seule, un atlas, mais (pour $r > 0$) la structure différentiable correspondante est différente de la structure standard, puisque les cartes $(\mathbb{R}, \mathrm{id})$ et $(\mathbb{R}, h_0)$ ne sont pas compatibles (le changement de cartes $h_0 \circ (\mathrm{id})^{-1} = h_0$ est différentiable et bijectif, mais l'application réciproque $h_0^{-1} : t \mapsto t^{1/3}$ n'est pas différentiable pour $t = 0$).

Exemple 4. Pour toute paramétrisation $\gamma : I \to \mathcal{A}$, $I =]a, b[$, d'un arc régulier simple ouvert $\mathcal{L}$, le couple $(\mathcal{L}, \gamma^{-1})$, où $\gamma^{-1} : \mathcal{L} \to I$ est l'application réciproque de $\gamma : I \to \mathcal{L}$, est une carte sur $\mathcal{L}$, et en outre d'après la proposition 1 de la leçon 1, deux telles cartes quelconques sont compatibles. Ceci exprime que *tout arc simple régulier ouvert est naturellement une variété à une dimension.*

Exemple 5. De façon analogue, toute surface élémentaire $\mathcal{X}$ est une variété différentiable à deux dimensions munie des cartes $(\mathcal{X}, \gamma^{-1})$, où γ est une paramétrisation arbitraire de $\mathcal{X}$.

Exemple 6. Une variété élémentaire non recouvrable par une seule carte est le cercle

$$\mathbf{S}^1 : x^2 + y^2 = 1.$$

Soient p_0 et q_0 les points $(-1, 0)$ et $(1, 0)$ de ce cercle et soient $U = \mathbf{S}^1 \setminus \{p_0\}$ et $V = \mathbf{S}^1 \setminus \{q_0\}$. Pour tout point $p \in U$ (resp. $q \in V$) désignons par $h(p)$ (resp. $k(p)$) l'angle $(\in]-\pi, \pi[)$ formé par le rayon vecteur de ce point et l'axe positif (resp. négatif) des abscisses. Il est clair que les applications obtenues

$$h : U \to]-\pi, \pi[, \qquad k : V \to]-\pi, \pi[$$

sont bijectives, i.e. les couples (U, h) et (V, k) sont des cartes sur $\mathbf{S}^1$ (avec $n = 1$). De plus, $U \cap V = \mathbf{S}^1 \setminus \{p_0, q_0\}$, les ensembles $h(U \cap V) =]-\pi, 0[\cup]0, \pi[$ et $k(U \cap V) =]-\pi, 0[\cup]0, \pi[$ (qui sont confondus) sont ouverts dans $\mathbb{R}$, et l'application

$$k \circ h^{-1} : h(U \cap V) \to k(U \cap V)$$

est définie par les formules

$$(k \circ h^{-1})(t) = \begin{cases} t + \pi & \text{si } t \in]-\pi, 0[, \\ t - \pi & \text{si } t \in]0, \pi[, \end{cases}$$

donc, est un difféomorphisme.

Les cartes (U, h) et (V, k) sont donc compatibles, et comme $U \cup V = \mathbf{S}^1$, elles forment un atlas et définissent par conséquent une structure différentiable (de classe C^ω) sur $\mathbf{S}^1$.

Exemple 7. On démontre de façon analogue que tout arc régulier simple fermé (sans extrémités) de l'espace affine $\mathcal{A}$ est une variété à une dimension, recouverte par deux cartes.

Exemple 8. Soient $U^{(+)}$, $U^{(-)}$, $V^{(+)}$, $V^{(-)}$ des sous-ensembles du cercle S^1 (des *demi-cercles ouverts*) composés de points $p = (x, y)$ pour lesquels respectivement $y > 0$, $y < 0$, $x > 0$, $x < 0$. Supposons ensuite que $h^{(+)}$, $h^{(-)}$, $k^{(+)}$, $k^{(-)}$ sont des applications de ces ensembles dans $\mathbb{R}$ définies par les formules $(x, y) \mapsto x$, $(x, y) \mapsto x$, $(x, y) \mapsto y$, $(x, y) \mapsto y$ (les restrictions des projections sur les axes de coordonnées). Chacune de ces applications applique bijectivement l'ensemble correspondant sur l'intervalle ouvert $]-1, 1[$ de l'axe $\mathbb{R}$, donc les couples

$$(U^{(+)}, h^{(+)}), (U^{(-)}, h^{(-)}), (V^{(+)}, k^{(+)}), (V^{(-)}, k^{(-)})$$

sont des cartes sur S^1 (de plus, x est coordonnée locale sur les cartes $U^{(\pm)}$, et y sur $V^{(\pm)}$). Il est aisé de voir que ces cartes sont compatibles (soit elles sont disjointes deux à deux, soit leur intersection est un quart de cercle se projetant sur les intervalles ouverts $]0, 1[$ ou $]-1, 0[$ de l'axe $\mathbb{R}$, les changements de cartes correspondants étant définis par les formules $y = \pm\sqrt{1 - x^2}$ et sont donc des difféomorphismes) et recouvrent le cercle S^1 tout entier, i.e. forment un atlas sur S^1. *Cet atlas est équivalent à celui de l'exemple* 6 (donc, il définit la même structure différentiable sur S^1). En effet, pour les cartes (U, h) et $(U^{(+)}, h^{(+)})$ par exemple, l'intersection $U \cap U^{(+)}$ est confondue avec $U^{(+)}$, et de plus $h(U^{(+)}) =]0, \pi[$, $h^{(+)}(U^{(+)}) =]-1, 1[$ et l'application

$$h^{(+)} \circ h^{-1} :]0, \pi[\to]-1, 1[$$

est définie par la formule $x = \cos t$, $0 < t < \pi$, donc est un difféomorphisme (la situation est la même pour les autres couples de cartes).

Exemple 9 (généralisation de l'exemple 8 à une dimension arbitraire). De façon analogue, sur la sphère unité à $n \geqslant 1$ dimensions

$$S^n : x_0^2 + x_1^2 + \ldots + x_n^2 = 1$$

de l'espace $\mathbb{R}^{n+1}$, les couples $(U_i^{(+)}, h_i^{(+)})$ et $(U_i^{(-)}, h_i^{(-)})$, $i = 0, 1, \ldots, n$, où $U_i^{(+)}$ et $U_i^{(-)}$ sont des *demi-sphères* définies respectivement par $x_i > 0$ et $x_i < 0$, et $h_i^{(+)}$, $h_i^{(-)}$, les restrictions à ces demi-sphères des projections sur le plan de coordonnées $x_i = 0$, sont des cartes deux à deux compatibles. Par exemple, l'application $h_i^{(+)}$ de la demi-sphère $U_i^{(+)}$ sur la boule $\mathring{B}^n : t_1^2 + \ldots + t_n^2 < 1$ de l'espace $\mathbb{R}^n$ (muni des coordonnées $t_1, \ldots, t_n$) est définie par les formules

$$t_a = \begin{cases} x_{a-1} & \text{si } a \leqslant i, \\ x_a & \text{si } a > i, \end{cases} \quad a = 1, \ldots, n$$

(de sorte que les nombres $t_1, \ldots, t_n$, i.e. les nombres $x_0, \ldots$ $\ldots, x_{i-1}, x_{i+1}, \ldots, x_n$, sont des coordonnées locales de la carte $(U_i^{(+)}, h_i^{(+)})$), et l'application réciproque $(h_i^{(+)})^{-1}: \mathring{\mathbb{B}}^n \to U_i^{(+)}$ par les formules

$$x_b = \begin{cases} t_{b+1} & \text{si } b < i, \\ \sqrt{1-\mathrm{t}^2} & \text{si } b = i, \quad b = 0, 1, \ldots, n, \\ t_b & \text{si } b > i, \end{cases}$$

où $\mathrm{t}^2 = t_1^2 + \ldots + t_n^2$. Donc, pour $i \leqslant j$, l'application $h_j^{(+)} \circ (h_{i-1}^{(+)})^{-1}$ sera définie sur la demi-boule $\mathring{\mathbb{B}}^n_{j>0} = \{\mathrm{t} \in \mathring{\mathbb{B}}^n;\ t_j > 0\}$ qu'elle appliquera sur la demi-boule $\mathring{\mathbb{B}}^n_{i>0} = \{\mathrm{t} \in \mathring{\mathbb{B}}^n;\ t_i > 0\}$ à l'aide des formules

$$(11) \qquad t'_a = \begin{cases} t_a & \text{si } a < i. \\ \sqrt{1-\mathrm{t}^2} & \text{si } a = i, \\ t_{a-1} & \text{si } i \leqslant a \leqslant j, \\ t_a & \text{si } a > j, \end{cases} \quad a = 1, \ldots, n.$$

Ces formules définissant visiblement un difféomorphisme, les cartes $(U_{i-1}^{(+)}, h_{i-1}^{(+)})$ et $(U_j^{(+)}, h_j^{(+)})$ sont donc compatibles. Pour les autres couples de cartes, dans les formules (11) les indices peuvent se déplacer et le signe moins peut apparaître devant le radical, mais ceci n'affecte évidemment pas la conclusion tirée sur le difféomorphisme.

Puisque les cartes construites recouvrent de toute évidence la sphère $\mathbb{S}^n$ tout entière, elles forment un atlas et structurent par conséquent la sphère $\mathbb{S}^n$ en variété différentiable.

Exemple 10. L'atlas construit dans l'exemple 9 est composé de $2n + 2$ cartes. Il s'avère que la structure différentiable sur $\mathbb{S}^n$ peut être définie par un atlas composé en tout de deux cartes (cf. exemple 6). A cet effet, pour tout point $p = (x_0, x_1, \ldots, x_n)$ de $\mathbb{S}^n$ distinct du point $p_0 = (1, 0, \ldots, 0)$ (le *pôle nord* de $\mathbb{S}^n$) considérons la droite p_0p de l'espace $\mathbb{R}^{n+1}$. Les équations canoniques de cette droite sont

$$\frac{X_0 - 1}{x_0 - 1} = \frac{X_1}{x_1} = \cdots = \frac{X_n}{x_n},$$

donc, elle coupe l'hyperplan $X_0 = 0$ (l'*hyperplan équatorial* de $\mathbb{S}^n$) au point de coordonnées

$$(12) \qquad t_1 = \frac{x_1}{1-x_0}, \ \ldots, \ t_n = \frac{x_n}{1-x_0}$$

(les coordonnées de l'hyperplan équatorial seront désignées par $t_1, \ldots, t_n$). Le point (12) s'appelle *projection stéréographique* du point $p \in \mathbb{S}^n \setminus \{p_0\}$ (comparer avec la leçon I.28). En désignant ce

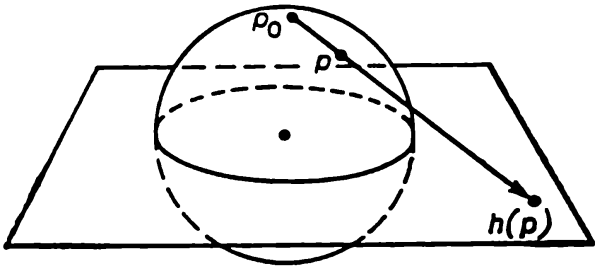
point par $h(p)$, on obtient une application h (qui est visiblement bijective) de l'ensemble $U = S^n \setminus \{p_0\}$ sur l'espace $\mathbb{R}^n$, i.e. une carte (U, h) (avec $h(U) = \mathbb{R}^n$).

De façon analogue, en partant du *pôle sud* $q_0 = (-1, 0, \ldots, 0)$ de la sphère S^n, on peut construire une projection stéréographique $k : V \to \mathbb{R}^n$, où $V = S^n \setminus \{q_0\}$, associant à un point arbitraire $p = (x_0, x_1, \ldots, x_n)$ de S^n, distinct de q_0, le point du plan équatorial de coordonnées

$$t'_1 = \frac{x_1}{1+x_0}, \ \ldots, \ t'_n = \frac{x_n}{1+x_0} \tag{13}$$

et donc obtenir une carte (V, k).

Pour les cartes construites, l'intersection $W = U \cap V$ est de la forme $S^n \setminus \{p_0, q_0\}$ et chaque application h et k donne de cette inter-



Projection stéréographique

section un ensemble ouvert $\mathbb{R}^n_* = \mathbb{R}^n \setminus \{0\}$ de l'espace $\mathbb{R}^n$. Le changement de cartes correspondant

$$k \circ h^{-1} : \mathbb{R}^n_* \to \mathbb{R}^n_* \tag{14}$$

sera défini par les formules obtenues en éliminant $x_0, x_1, \ldots, x_n$ des formules (12) et (13). Mais en vertu des formules (12), on a $x_i = (1 - x_0)\, t_i$, donc,

$$1 = x_0^2 + x_1^2 + \ldots + x_n^2 = x_0^2 + (1-x_0)^2 (t_1^2 + \ldots + t_n^2) = \\ = x_0^2 + (1-x_0)^2 t^2,$$

i.e.

$$(1+t^2)\, x_0^2 - 2t^2 x_0 + t^2 - 1 = 0.$$

Cette équation a deux racines : une racine $x_0 = 1$ qui est inutile (car elle correspond au point exclu p_0) et une racine

$$x_0 = \frac{t^2-1}{1+t^2}$$

pour laquelle on a les formules

$$1-x_0=\frac{2}{1+t^2}, \quad 1+x_0=\frac{2t^2}{1+t^2}.$$

Donc, pour tout $i=1, \ldots, n$,

$$x_i=(1-x_0)\,t_i=\frac{2t_i}{1+t^2}$$

et, par suite,

$$t'_i=\frac{x_i}{1+x_0}=\frac{t_i}{t^2}.$$

Ceci prouve que l'application (14) est définie par les formules

$$t'_1=\frac{t_1}{t^2}, \quad \ldots, \quad t'_n=\frac{t_n}{t^2}$$

et, par conséquent, est un difféomorphisme.

Donc, les cartes (U, h) et (V, k) sont compatibles et forment un atlas.

Exercice 2. Montrer que cet atlas est équivalent à celui de l'exemple 9 et définit donc la même structure différentiable sur la sphère $\mathbf{S}^n$.

Exemple 11. On rappelle que l'*espace projectif réel* $\mathbb{R}P^n$ *à n dimensions* est l'ensemble de toutes les classes $(x_0 : x_1 : \ldots : x_n)$ des vecteurs proportionnels $(x_0, x_1, \ldots, x_n) \neq 0$ ou, ce qui est équivalent, l'ensemble de toutes les droites de l'espace $\mathbb{R}^{n+1}$ passant par le point 0. Soit U_i, $i = 0, \ldots, n$, l'ensemble de tous les points $p = (x_0 : x_1 : \ldots : x_n)$ de l'espace $\mathbb{R}P^n$ pour lesquels $x_i \neq 0$ et soit $h_i : U_i \to \mathbb{R}^n$ une application associant à tout point $p = (x_0 : x_1 : \ldots : x_n)$ de U_i le point $(t_1, \ldots, t_n)$ de $\mathbb{R}^n$ pour lequel

$$t_a=\begin{cases}\dfrac{x_{a-1}}{x_i} & \text{si } a \leqslant i,\\[2mm] \dfrac{x_a}{x_i} & \text{si } a > i,\end{cases} \qquad a=1, \ldots, n. \tag{15}$$

Il est clair que l'application h_i est bijective, de sorte que le couple (U_i, h_i), $i = 0, 1, \ldots, n$, est une carte sur $\mathbb{R}P^n$.

Pour tous indices i, $j \neq i$, l'intersection $U_i \cap U_j$ est composée des points $p = (x_0 : x_1 : \ldots : x_n)$ pour lesquels $x_i \neq 0$ et $x_j \neq 0$, et les ensembles $h_i(U_i \cap U_j)$ et $h_j(U_i \cap U_j)$ sont composés (pour $j > i$) des points $\mathbf{t} = (t_1, \ldots, t_n)$ de $\mathbb{R}^n$ pour lesquels respectivement $t_j \neq 0$ et $t_{i+1} \neq 0$. Il est clair que ces ensembles sont ouverts.

Puisque l'application h_i opère d'après les formules (15), et l'application h_j, d'après des formules analogues obtenues en remplaçant i par j, le changement de cartes $h_j \circ h_i^{-1}$ transforme tout point $(t_1, \ldots, t_n) \in h_i(U_i \cap U_j)$ en un point $(t'_1, \ldots, t'_n) \in h_j(U_i \cap U_j)$ tel que

$$t'_a = \begin{cases} \dfrac{t_a}{t_j} & \text{si } a \leqslant i \text{ ou } a > j, \\ \dfrac{1}{t_j} & \text{si } a = i+1, \\ \dfrac{t_{a-1}}{t_j} & \text{si } i+1 < a \leqslant j, \end{cases} \qquad a = 1, \ldots, n,$$

donc, est un difféomorphisme.

Ceci prouve que les cartes (U_i, h_i) et (U_j, h_j) sont deux à deux compatibles et, par suite, ces cartes définissent une structure différentiable sur $\mathbb{R}P^n$.

Remarquons que dans tous ces exemples (hormis bien sûr les exemples 4, 5 et 7) on a obtenu des variétés $\mathbb{R}$-analytiques (de classe C^ω).

LEÇON 7

Topologie d'une variété différentiable. — Sous-variétés ouvertes. — Points intérieurs et voisinages. — Homéomorphismes. — Premier axiome de dénombrabilité et euclidicité locale. — Deuxième axiome de dénombrabilité. — Variétés non séparées. — Structures différentiables sur un espace topologique. — Variétés topologiques. — Variétés 0-dimensionnelles. — Catégorie TOP. — Catégorie DIFF. — Transport de différentiabilité.

Soit $\mathscr{X}$ une variété différentiable (ou topologique) arbitraire.

Définition 1. On dit qu'un sous-ensemble $O \subset \mathscr{X}$ est *ouvert* (dans $\mathscr{X}$) si pour toute carte (U, h) de la variété $\mathscr{X}$ l'ensemble $h(O \cap U) \subset \mathbb{R}^n$ est ouvert (dans $\mathbb{R}^n$).

Pour $\mathscr{X} = \mathbb{R}^n$, cette définition est de toute évidence celle d'un ordinaire ensemble ouvert dans $\mathbb{R}^n$.

On sait du cours d'analyse que les ensembles ouverts dans l'espace $\mathscr{X} = \mathbb{R}^n$ (ainsi que dans tout espace métrique $\mathscr{X}$) jouissent des trois propriétés suivantes:

1° l'ensemble vide $\varnothing$ et l'espace $\mathscr{X}$ sont ouverts;

2° la réunion d'une famille quelconque d'ensembles ouverts est un ensemble ouvert;

3° l'intersection d'une famille finie quelconque d'ensembles ouverts est un ensemble ouvert.

Puisque

$$\bigcup_\alpha h(O_\alpha \cap U) = h\left(\left(\bigcup_\alpha O_\alpha\right) \cap U\right)$$

et

$$\bigcap_\alpha h(O_\alpha \cap U) = h\left(\left(\bigcap_\alpha O_\alpha\right) \cap U\right)$$

pour toute carte (U, h) et toute famille $\{O_\alpha\}$ de sous-ensembles de la variété $\mathscr{X}$, ces propriétés caractérisent aussi les ensembles ouverts dans $\mathscr{X}$.

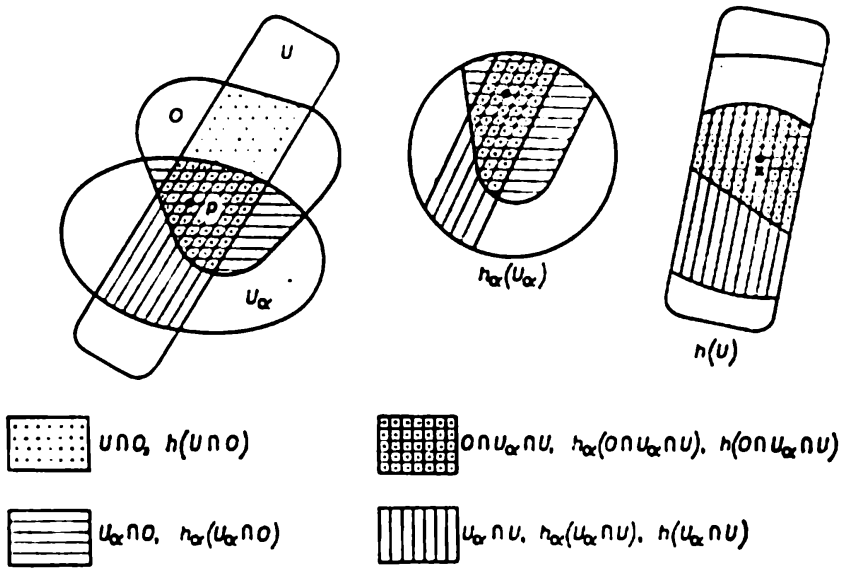
Définition 2. Un ensemble $\mathbb{T}$ de sous-ensembles d'un ensemble $\mathscr{X}$ s'appelle *topologie* (ou *structure topologique*) sur $\mathscr{X}$, et les ensembles de $\mathbb{T}$, *ensembles ouverts* si ces ensembles possèdent les propriétés 1° à 3°. L'ensemble $\mathscr{X}$ muni de la topologie $\mathbb{T}$ (ou, à proprement parler, le couple $(\mathscr{X}, \mathbb{T})$) s'appelle *espace topologique*.

Nous pouvons donc dire que la définition 1 munit chaque variété d'une certaine topologie ou, en d'autres termes, qu'en vertu de cette définition *toute variété est un espace topologique*. En ce sens, la topologie sur $\mathscr{X}$ est une structure dérivée de la structure différentiable de $\mathscr{X}$.

Cette situation est identique à celle des espaces métriques dont la topologie est une structure dérivée de la métrique.

Dans la suite, en parlant d'une topologie sur une variété différentiable (ou topologique), on aura toujours en vue la topologie de la définition 1.

En vertu de la définition 1, pour s'assurer qu'un sous-ensemble $O \subset \mathscr{X}$ est ouvert ou non, il faut envisager les ensembles $h\,(O \cap U)$



pour toutes les cartes (U, h) de la variété $\mathscr{X}$. Ceci est bien sûr impossible et le problème qui se pose est de réduire le nombre de ces cartes. Il s'avère qu'il suffit d'envisager les cartes d'un seul atlas et plus même les seules cartes recouvrant l'ensemble O. Plus exactement, on a la proposition suivante:

Proposition 1. *Soit* $\{(U_\alpha, h_\alpha)\}$ *une famille de cartes de la variété* $\mathscr{X}$, *telle que*

$$O \subset \bigcup_\alpha U_\alpha.$$

Pour que l'ensemble O *soit ouvert dans* $\mathscr{X}$, *il suffit que pour tout* α *l'ensemble* $h_\alpha\,(O \cap U_\alpha)$ *le soit dans* $\mathbb{R}^n$.

Démonstration. Il faut prouver que si l'ensemble O satisfait aux conditions de cette proposition, alors pour toute carte (U, h) de la variété $\mathscr{X}$ l'ensemble $h\,(O \cap U)$ est ouvert dans $\mathbb{R}^n$, i.e. tout point $\mathbf{x}$ de cet ensemble, s'il n'est pas vide, est un point intérieur. Considérons le point $p = h^{-1}(\mathbf{x})$ de la variété $\mathscr{X}$.

Puisque $\mathbf{x} \in h\,(O \cap U)$, alors $p \in O \cap U$ et en particulier $p \in O$. Il existe donc un α tel que $p \in U_\alpha$.

Les cartes (U_α, h_α) et (U, h) étant compatibles, les ensembles $h_\alpha\,(U_\alpha \cap U)$ et $h\,(U_\alpha \cap U)$ sont ouverts dans $\mathbb{R}^n$. D'autre part,

l'ensemble $h_\alpha (O \cap U_\alpha)$ est ouvert aussi par hypothèse. Donc, il en est de même de l'ensemble

$$h_\alpha (O \cap U_\alpha \cap U) = h_\alpha (O \cap U_\alpha) \cap h_\alpha (U_\alpha \cap U).$$

Par ailleurs, l'image de cet ensemble par l'homéomorphisme

$$h \circ h_\alpha^{-1} : h_\alpha(U_\alpha \cap U) \to (U_\alpha \cap U)$$

est l'ensemble $h (O \cap U_\alpha \cap U)$ qui est ouvert aussi. Comme

$$\mathbf{x} \in h (O \cap U_\alpha \cap U) \subset h (O \cap U),$$

ceci prouve que le point $\mathbf{x}$ est un point intérieur de l'ensemble $h (O \cap U)$. □

Corollaire 1. *Pour toute carte (U, h) de la variété $\mathscr{X}$, un sous-ensemble $V \subset U$ est ouvert dans $\mathscr{X}$ si et seulement si l'ensemble $h (V)$ l'est dans $\mathbb{R}^n$.*

En particulier, l'ensemble U est ouvert dans $\mathscr{X}$. □

* * *

Pour tout sous-ensemble ouvert O d'une variété $\mathscr{X}$ (de classe C^r), toutes les cartes (U, h) telles que $U \subset O$ forment de toute évidence un atlas sur O (cf. exemple 2 de la leçon précédente). Cet atlas est maximal, i.e. est une structure différentiable sur O (de même classe C^r).On dit de cette structure différentiable qu'elle est *induite* par celle de la variété $\mathscr{X}$, et l'ensem le O muni de cette structure s'appelle *sous-variété ouverte* de la variété $\mathscr{X}$. Par définition, $\dim O = \dim \mathscr{X}$.

Dans la suite, tout ensemble ouvert O d'une variété $\mathscr{X}$ sera traité comme une variété munie de la structure différentiable induite.

Remarquons que chaque atlas $\mathsf{A} = \{(U_\alpha, h_\alpha)\}$ de la variété $\mathscr{X}$ définit un atlas $O \cap \mathsf{A}$ de la variété O, constitué de cartes $(O \cap U_\alpha, h_\alpha \mid_{O \cap U_\alpha})$. Donc, pour obtenir une structure différentiable induite sur O, il suffit d'envisager les cartes d'un seul atlas.

* * *

De la propriété 2° des ensembles ouverts il s'ensuit immédiatement que pour tout sous-ensemble A d'un espace topologique $\mathscr{X}$ il existe un ensemble ouvert maximal (éventuellement vide) contenu dans A (ce sera la réunion de tous les ensembles ouverts contenus dans A). Cet ensemble ouvert est désigné par $\operatorname{Int} A$ (ou $\mathring{A}$) et s'appelle *intérieur* de A. Ses points s'appellent *points intérieurs* de A.

Il est clair qu'*un ensemble A est ouvert si et seulement si* $\operatorname{Int} A = A$, *i.e. si tous ses points sont intérieurs.*

Tout ensemble ouvert contenant un point (ou, de façon plus générale, un sous-ensemble) d'un espace topologique $\mathscr{X}$ s'appelle *voisi-*

nage de ce point (de ce sous-ensemble). Par définition, *un point est un point intérieur à un sous-ensemble A si et seulement si A contient un voisinage de ce point.*

Donc, pour $\mathscr{X} = \mathbb{R}^n$, la notion de point intérieur introduite est confondue avec celle qui est connue en analyse,

Le support U de toute carte (U, h) d'une variété $\mathscr{X}$ est un voisinage de tout point $p \in U$. Pour cette raison, les supports des cartes de $\mathscr{X}$ sont appelés aussi *voisinages de coordonnées.*

Il est clair que tout voisinage V d'un point $p \in U$, contenu dans un voisinage de coordonnées U, est aussi un voisinage de coordonnées (avec l'application $h \mid_V$). Donc, *tout voisinage O d'un point p contient un voisinage de coordonnées* (un tel voisinage sera par exemple l'intersection $O \cap U$).

Un ensemble de voisinages d'un point p d'un espace topologique $\mathscr{X}$ s'appelle *base de voisinages* (ou *système fondamental de voisinages*) si tout voisinage de p contient un voisinage de cet ensem le. (De façon imagée, la base doit contenir des voisinages aussi petits que l'on veut.)

Dans cette terminologie, la proposition prouvée exprime que *les voisinages de coordonnées de tout point d'une variété forment une base de voisinages.*

* * *

On dit qu'une application $f : \mathscr{X} \to \mathscr{Y}$ d'espaces topologiques est un *homéomorphisme* si elle est bijective et un ensemble $O \subset \mathscr{X}$ est ouvert dans $\mathscr{X}$ si et seulement si l'ensemble $fO \subset \mathscr{Y}$ l'est dans $\mathscr{Y}$. Donc, un homéomorphisme est une application bijective établissant une correspondance biunivoque entre les topologies.

Des espaces $\mathscr{X}$ et $\mathscr{Y}$ sont dits *homéomorphes* s'il existe au moins un homéomorphisme de $\mathscr{X}$ sur $\mathscr{Y}$. Il est clair que les propriétés topologiques (i.e. les propriétés formulées exclusivement en termes d'ensembles ouverts) des espaces homéomorphes sont identiques. Donc, de ce point de vue il n'y a pas de différence entre les espaces homéomorphes.

* * *

Un exemple de propriété topologique est la propriété d'un espace topologique $\mathscr{X}$ de posséder en tout point p une base dénombrable de voisinages. On dit de tels espaces qu'ils satisfont au *premier axiome de dénombrabilité* (ou encore que ce sont des espaces de *poids local dénombrable*).

Il est aisé de voir que *tout espace métrique $\mathscr{X}$ satisfait au premier axiome de dénombrabilité* (les boules ouvertes centrées en un point quelconque de $\mathscr{X}$ et de rayons rationnels forment une base dénombra-

ble de voisinages). En particulier, *l'espace* $\mathbb{R}^n$ *satisfait au premier axiome de dénombrabilité.*

Tout sous-ensemble A d'un espace topologique $\mathcal{X}$ est muni d'une topologie (dite *topologie induite*) dont les ensembles ouverts sont les intersections $O \cap A$ des ouverts $O \subset \mathcal{X}$ et de A. Un sous-ensemble A muni de cette topologie s'appelle *sous-espace* de l'espace $\mathcal{X}$.

Il est clair que tout sous-espace d'un espace satisfaisant au premier axiome de dénombrabilité satisfait aussi à cet axiome. Donc, en particulier, *tout sous-espace de l'espace* $\mathbb{R}^n$ *satisfait au premier axiome de dénombrabilité.*

On dit qu'une famille $\{U_\alpha\}$ d'ensembles ouverts d'un espace topologique $\mathcal{X}$ est un *recouvrement ouvert* de $\mathcal{X}$ si tout p int de $\mathcal{X}$ appartient au moins à un élément de cette famille, i.e. si

$$\bigcup_\alpha U_\alpha = \mathcal{X}.$$

Par exemple, les supports U_α des cartes (U_α, h_α) d'un atlas arbitraire d'une variété $\mathcal{X}$ forment un recouvrement ouvert de $\mathcal{X}$.

Un espace $\mathcal{X}$ est dit *localement euclidien* s'il possède un recouvrement ouvert $\{U_\alpha\}$ dont chaque élément U_α est homéomorphe à un ensemble ouvert de l'espace $\mathbb{R}^n$ (donc, satisfait au premier axiome de dénombrabilité).

Il est clair que si un espace topologique $\mathcal{X}$ possède un recouvrement ouvert $\{U_\alpha\}$ dont tous les éléments U_α satisfont (pour la topologie induite) au premier axiome de dénombrabilité, cet espace satisfera lui-même à cet axiome. (On exprime ce fait en disant que la propriété de vérifier le premier axiome de dénombrabilité est une *propriété locale.*) Donc, *tout espace localement euclidien satisfait au premier axiome de dénombrabilité.*

Par ailleurs, du corollaire 1 il s'ensuit immédiatement que *pour toute carte* (U, h) *d'une variété* $\mathcal{X}$ *l'application* $h: U \to h(U)$ *est un homéomorphisme* (et de plus l'ensemble $h(U)$ est ouvert dans $\mathbb{R}^n$).

[Pour cette raison, les applications h traitées comme des applications sur $h(U)$ s'appellent *homéomorphismes de coordonnées.*]

Donc, puisque les voisinages de coordonnées forment un recouvrement ouvert de $\mathcal{X}$ (cette propriété caractérise même les voisinages de coordonnées qui sont des domaines de définition de cartes d'un atlas), *toute variété est un espace localement euclidien qui satisfait par conséquent au premier axiome de dénombrabilité.*

* * *

On dit qu'un ensemle $\mathcal{B}$ d'ensembles ouverts d'un espace topologique $\mathcal{X}$ est une *base* (ou, de façon plus usuelle, une *base d'ensembles ouverts*) de cet espace si tout ensemble ouvert de $\mathcal{X}$ est la réunion

d'ensembles de $\mathfrak{B}$. A propos des espaces possédant une *base dénombrable* (i.e. une base composée d'une quantité dénombrable d'ensembles ouverts) on dit qu'ils satisfont au *deuxième axiome de dénombrabilité* (et aussi que ce sont des espaces de *poids dénombrable*).

Il est clair que tout espace satisfaisant au deuxième axiome de dénombrabilité satisfait au premier, mais la réciproque n'est généralement pas vraie.

Dans cette terminologie, le lemme 2 de la leçon 1 affirme (s'agissant de l'espace $\mathbb{R}^n$) que toutes les boules rationnelles ouvertes de $\mathbb{R}^n$, i.e. les boules de rayons rationnels et dont les centres possèdent des coordonnées rationnelles, forment une base de cet espace. Donc, *l'espace $\mathbb{R}^n$ satisfait au deuxième axiome de dénombrabilité.*

Il existe toutefois des variétés ne satisfaisant pas au deuxième axiome de dénombrabilité (de sorte que sur ce plan le deuxième axiome contraste avec le premier). Un exemple élémentaire est la réunion disjonctive d'une famille dénombrable d'espaces $\mathbb{R}^n$. (Nous citerons un exemple plus intéressant à la leçon 11.)

En munissant un ensemble $\mathscr{X}$ d'une topologie, souvent on n'indique explicitement qu'une base de cet ensemble. Ceci étant, il est évident qu'*un ensemble $\mathfrak{B}$ de parties d'un ensemble $\mathscr{X}$ est une base d'ensembles ouverts pour une topologie de $\mathscr{X}$ si et seulement si l'intersection de deux éléments quelconques de $\mathfrak{B}$ est la réunion d'éléments de $\mathfrak{B}$.*

En particulier, tout ensemble de parties stable pour les intersections est une base pour une certaine topologie.

* * *

On dit qu'un espace topologique $\mathscr{X}$ est *séparé* (ou *de Hausdorff*) si deux quelconques de ses points p et q possèdent des voisinages disjoints.

Il est clair que tout espace métrique est séparé, donc en particulier l'espace $\mathbb{R}^n$ est séparé. *Il existe* néanmoins *des variétés différentiables non séparées.*

Exemple 1. Soit $\mathscr{X} = (\mathbb{R} \setminus \{0\}) \cup \{p_0, q_0\}$, où p_0, q_0 sont des éléments (n'appartenant pas à $\mathbb{R} \setminus \{0\}$) et soient

$$U = \mathscr{X} \setminus \{q_0\} = (\mathbb{R} \setminus \{0\}) \cup \{p_0\},$$
$$V = \mathscr{X} \setminus \{p_0) = (\mathbb{R} \setminus \{0\}) \cup \{q_0\}.$$

Définissons les applications bijectives

$$h : U \to \mathbb{R}, \quad k : V \to \mathbb{R}$$

par les formules

$$h(p) = \begin{cases} p & \text{si } p \in \mathbb{R} \setminus \{0\}, \\ 0 & \text{si } p = p_0 \end{cases}$$

pour tout $p \in U$, et

$$k(q) = \begin{cases} q & \text{si } q \in \mathbb{R} \setminus \{0\}, \\ 0 & \text{si } q = q_0 \end{cases}$$

pour tout $q \in V$.

Il est clair que les cartes (U, h) et (V, k) forment un atlas et donc structurent $\mathscr{X}$ en une variété différentiable (de classe C^{ω}). Deux voisinages quelconques des points p_0 et q_0 se coupent pour la topologie de cette variété, donc cette topologie n'est pas séparée.

La variété différentiable $\mathscr{X}$ construite s'appelle *droite dédoublée en* 0 (ou encore *droite non séparée avec un point singulier double* 0).

On définit de façon analogue une *droite non séparée avec un point singulier* 0 *de multiplicité n*, où *n est un entier* $\geqslant 2$.

* * *

Très souvent on a à structurer en variété différentiable un ensemble $\mathscr{X}$ muni déjà d'une topologie, i.e. un espace topologique.

Dans ce cas on admet toujours tacitement que cette structure doit définir la topologie dont est muni l'ensemble $\mathscr{X}$, i.e. comme on dit généralement, cette structure doit être *compatible* avec la topologie.

Pour cela il faut de toute évidence que l'atlas qui définit sur $\mathscr{X}$ la structure différentiable possédant cette propriété soit composé de cartes (U, h) dont les supports U soient ouverts et les applications $h : U \to h(U)$ soient des homéomorphismes. Il se trouve que cette condition est suffisante.

Proposition 2. *Soit* $\mathsf{A} = \{(U_\alpha, h_\alpha)\}$ *un atlas d'un espace topologique* $\mathscr{X}$, *tel que pour tout* α *l'ensemble* U_α *est ouvert dans* $\mathscr{X}$ *et l'application* $h_\alpha : U_\alpha \to h_\alpha(U_\alpha)$ *est un homéomorphisme. La structure différentiable définie par l'atlas* A *est alors compatible avec la topologie de l'espace* $\mathscr{X}$.

Démonstration. En vertu de la proposition 1, un ensemble $O \subset \mathscr{X}$ est ouvert pour la topologie $\mathbb{T}_{\mathsf{A}}$ définie par la structure différentiable de l'atlas A si et seulement si l'ensemble $h_\alpha(O \cap U_\alpha)$ est ouvert pour tout α (dans $\mathbb{R}^n$, donc dans $h_\alpha(U_\alpha)$), i.e. puisque les applications h_α sont des homéomorphismes, si et seulement si l'ensemble $O \cap U_\alpha$ est ouvert (dans U_α, donc dans $\mathscr{X}$, puisque U_α est ouvert dans $\mathscr{X}$).

D'autre part, si un ensemble $O \subset \mathscr{X}$ est ouvert dans $\mathscr{X}$, il en est de même de $O \cap U_\alpha$ (puisque U_α est ouvert dans $\mathscr{X}$) et, si tous les ensembles $O \cap U_\alpha$ le sont dans $\mathscr{X}$, il en est de même de O en vertu de l'égalité

$$O = \bigcup_\alpha (O \cap U_\alpha). \tag{1}$$

Donc, les ensembles ouverts pour la topologie $\mathbb{T}_{\mathsf{A}}$ ne sont autres que les ensembles ouverts pour la topologie de l'espace $\mathscr{X}$. □

Remarque 1. Pour éclaircir la démonstration produite, il convient de signaler qu'elle se ramène en fait à deux remarques triviales. La première est que la topologie de chaque ensemble U_α est définie exclusivement par l'application h_α (un ensemble $O \subset U_\alpha$ est ouvert dans U_α si et seulement si l'ensemble $h_\alpha(O)$ l'est dans $\mathbb{R}^n$), et la deuxième que pour tout espace topologique $\mathscr{X}$ et tout recouvrement ouvert $\{U_\alpha\}$ de $\mathscr{X}$, la topologie de $\mathscr{X}$ est définie de façon unique (moyennant l'égalité (1)) par les topologies des éléments U_α de ce recouvrement.

Remarque 2. D'après la proposition 2, une variété différentiable peut être définie comme un espace topologique $\mathscr{X}$ sur lequel est donné un atlas dont les supports U_α des cartes (U_α, h_α) sont ouverts et les applications h_α sont des homéomorphismes. Ceci étant, la condition que les ensembles $h_\alpha(U_\alpha \cap U_\beta)$ et $h_\beta(U_\alpha \cap U_\beta)$ soient ouverts dans $\mathbb{R}$ pour tous α et β sera automatiquement remplie. C'est précisément cette définition que l'on rencontre habituellement (parfois sous une forme verbale légèrement différente) malgré un défaut méthodologique évident (une définition analogue des espaces métriques, mettant ce défaut plus en relief, dirait qu'un espace métrique est un espace topologique $\mathscr{X}$ pour lequel est définie une fonction continue $\rho : \mathscr{X} \times \mathscr{X} \to \mathbb{R}$ satisfaisant aux axiomes classiques de l'espace métrique).

Dans les exemples de la leçon précédente, nous avons précisément muni de structure différentiable des ensembles ($\mathbb{R}^n$, S^n et $\mathbb{R}P^n$) qui étaient déjà des espaces topologiques (et même métriques). [Pour distance des points de la sphère S^n on prend l'angle de deux rayons vecteurs et pour distance des points de l'espace $\mathbb{R}P^n$, l'angle de ces points traités comme des droites de l'espace $\mathbb{R}^{n+1}$. Dans les deux cas (vérifier-le !) les axiomes de l'espace métrique sont satisfaits.]

Exercice 1. Montrer que les structures différentiables définies sur les espaces $\mathbb{R}^n$, S^n et $\mathbb{R}P^n$ dans les exemples 1 à 8 de la leçon 6 sont compatibles avec les topologies de ces espaces.

Remarque 3. L'exemple 3 de la leçon 6 montre qu'un espace topologique peut être muni de plusieurs structures différentiables d i s t i n c t e s (de classe C^ω) compatibles avec sa topologie.

* * *

Pour $r = 0$, i.e. pour les variétés topologiques, la situation est différente.

Soit $\mathscr{X}$ une variété topologique de dimension n et soit (U, h) une carte sur $\mathscr{X}$. Par définition, cela signifie que $U \subset \mathscr{X}$ et que h est une application bijective de U sur $h(U)$, où $h(U)$ est un ouvert de $\mathbb{R}^n$. Ceci étant, comme déjà signalé, si (U, h) est une carte de la va-

riété $\mathcal{X}$, alors U est ouvert dans $\mathcal{X}$ et $h: U \to h(U)$ est un homéomorphisme. Il s'avère que la réciproque est vraie pour $r = 0$.

Proposition 3. *Une carte (U, h) d'une variété topologique $\mathcal{X}$, pour laquelle U est un ouvert de $\mathcal{X}$ et l'application $h: U \to h(U)$ est un homéomorphisme, est une carte de $\mathcal{X}$ (appartenant à son atlas maximal A_{max}).*

Pour prouver cette proposition on se servira du lemme suivant:

Lemme 1. *Si les supports U et V de cartes (U, h) et (V, k) d'un espace topologique $\mathcal{X}$ sont des ouverts de $\mathcal{X}$ et les applications $h: U \to h(U)$ et $k: V \to k(V)$ sont des homéomorphismes, ces cartes sont compatibles (dans la classe C^0).*

La proposition 3 est la conséquence directe du lemme 1, puisque d'après ce lemme la carte (U, h) de la proposition 3 est compatible avec toute carte (V, k) de l'atlas maximal A_{max} de la variété $\mathcal{X}$, donc appartient à cet atlas. Il nous faut par conséquent démontrer uniquement le lemme 1.

Démonstration du lemme 1. L'intersection $U \cap V$ est ouverte dans $\mathcal{X}$ (donc dans U). Les applications h et k étant des homéomorphismes, les ensembles $h(U \cap V)$ et $k(U \cap V)$ sont donc ouverts (respectivement dans $h(U)$ et $k(V)$ et par conséquent dans $\mathbb{R}^n$).

D'autre part, les applications :

$$h|_{U \cap V}: U \cap V \to h(U \cap V) \text{ et } k|_{U \cap V}: U \cap V \to k(U \cap V)$$

étant visiblement des homéomorphismes, il en est de même de leur composée

$$h \circ k^{-1}: k(U \cap V) \to h(U \cap V).$$

Donc, les cartes (U, h) et (V, k) sont compatibles. □

La proposition 3 exprime que pour $r = 0$ les cartes de la variété $\mathcal{X}$ sont caractérisées de façon purement topologique, i.e. *la structure d'une variété topologique sur un espace topologique $\mathcal{X}$ (si elle existe) est définie de façon unique.* En d'autres termes, la structure d'une variété topologique sur un espace topologique $\mathcal{X}$ n'apporte rien de nouveau à cet espace, et la classe des variétés topologiques est tout simplement une sous-classe de la classe de tous les espaces topologiques.

Pour dire si un espace topologique $\mathcal{X}$ est une variété topologique à n dimensions, il faut envisager tous les sous-ensembles ouverts $U \subset \mathcal{X}$ homéomorphes aux sous-ensembles ouverts de $\mathbb{R}^n$. L'espace $\mathcal{X}$ sera une variété si et seulement si toutes les cartes (U, h), où $h: U \to h(U)$ est un homéomorphisme (et $h(U)$ un ouvert de $\mathbb{R}^n$), forment un atlas, i.e. en raison de la compatibilité de deux telles cartes, si et seulement si les ensembles U forment un recouvrement ouvert de l'espace $\mathcal{X}$. Ceci prouve que *les variétés topologiques ne sont autres que les espaces localement euclidiens.*

Remarque 4. Il existe des espaces localement euclidiens (= variétés topologiques) qui ne peuvent être munis d'une structure de variété différentiable (de classe C^r avec $r > 0$). (De telles variétés sont dites *indifférentiables*.) Les conditions nécessaires et suffisantes d'existence d'une structure différentiable sur une variété topologique, compatible avec sa topologie, sont connues mais fort complexes.

Les variétés topologiques ne seront pas étudiées dans la suite si le contraire n'est pas spécifié.

* * *

Comme déjà dit à la leçon 6, le cas $n = 0$ ne sera pas exclu. En fait, il est peu intéressant.

En effet, pour $n = 0$ l'espace $\mathbb{R}^n$ est composé du seul point 0, donc, dans une variété 0-dimensionnelle $\mathscr{X}$, tout point possède un seul voisinage de coordonnées composé de ce point. Donc, dans une variété 0-dimensionnelle $\mathscr{X}$, tout point (par conséquent, tout sous-ensemble de $\mathscr{X}$) est un ensemble ouvert. Un espace topologique possédant cette propriété est dit *discret*. Puisque tout espace discret possède visiblement pour seule structure la structure d'une variété 0-dimensionnelle (pour laquelle tout point est le support d'une carte), on voit que *les variétés 0-dimensionnelles ne sont autres que les espaces discrets*.

Signalons que les cartes d'une variété 0-dimensionnelle sont disjointes. Nous avons donc le droit d'admettre que cette variété est de classe C^r quelconque.

Toute variété 0-dimensionnelle (= espace discret) $\mathscr{X}$ est métrisable au moyen d'une métrique pour laquelle la distance de deux points distincts quelconques est égale à l'unité. Ceci justifie la représentation de $\mathscr{X}$ comme un ensemble parsemé de points isolés distincts.

Dans la suite, le cas $n = 0$ sera en général écarté.

* * *

On dit qu'une application $f : \mathscr{X} \to \mathscr{Y}$ d'espaces topologiques est *continue en un point* $p \in \mathscr{X}$ si pour tout voisinage V du point $f(p)$ de $\mathscr{Y}$ il existe un voisinage U du point p dans $\mathscr{X}$ tel que $f(U) \subset V$ (comparer avec la (ε, δ)-définition d'une fonction continue).

Une application $f : \mathscr{X} \to \mathscr{Y}$ continue en chaque point $p \in \mathscr{X}$ est dite *continue sur* $\mathscr{X}$.

Il est aisé de voir qu'*une application* $f : \mathscr{X} \to \mathscr{Y}$ *est continue sur* $\mathscr{X}$ *si et seulement si l'image réciproque* $f^{-1} O = \{p \in \mathscr{X} ; f(p) \in O\}$ *de tout ensemble ouvert* $O \subset \mathscr{Y}$ *est un ensemble ouvert dans* $\mathscr{X}$. En effet, si une application f est continue, pour tout point $p \in f^{-1}O$ il existe (puisque O est un voisinage du point $f(p)$) un voisinage U dans $\mathscr{X}$ tel

que $fU \subset O$. Mais alors $U \subset f^{-1}O$ et, par suite, le point p est un point intérieur de l'ensemble $f^{-1}O$. Donc, l'ensemble $f^{-1}O$ est ouvert. Réciproquement, si pour tout ouvert $O \subset \mathcal{Y}$ l'ensemble $f^{-1}O \subset \mathcal{X}$ est ouvert, alors en particulier pour tout point $p \in \mathcal{X}$ et tout voisinage V du point $f(p)$ dans $\mathcal{Y}$ l'ensemble $U = f^{-1}V$ (contenant le point p) est ouvert (i.e. est un voisinage du point p). Puisque $fU \subset V$, on en déduit que l'application f est continue en p. □

Il s'ensuit que *les homéomorphismes $f : \mathcal{X} \to \mathcal{Y}$ sont exactement des applications bijectives continues dont les applications réciproques $f^{-1} : \mathcal{Y} \to \mathcal{X}$ sont aussi continues.*

Une application injective continue $f : \mathcal{X} \to \mathcal{Y}$ qui est un homéomorphisme sur le sous-espace $f\mathcal{X} \subset \mathcal{Y}$ s'appelle *monéomorphisme* (cf. leçon 1).

Remarquons que si un ensemble $O \subset \mathcal{X}$ est ouvert dans $\mathcal{X}$ et si l'application $f : \mathcal{X} \to \mathcal{Y}$ est continue, l'ensemble fO peut ne pas être ouvert dans $\mathcal{Y}$. Ce sera visiblement le cas si le sous-ensemble $f\mathcal{X}$ de $\mathcal{Y}$ ne possède aucun point intérieur. (De façon plus concrète: $\mathcal{X} = \mathbb{R}$, $\mathcal{Y} = \mathbb{R}^2$ et f est l'injection $x \mapsto (x, 0)$.)

Les applications continues $f : \mathcal{X} \to \mathcal{Y}$ qui sont telles que pour tout ensemble ouvert O dans $\mathcal{X}$ l'ensemble $fO \subset \mathcal{Y}$ est ouvert dans $\mathcal{Y}$ sont dites *ouvertes*.

Exercice 2. Montrer que:

a) Pour tout espace topologique $\mathcal{X}$ l'application identique $\mathrm{id} : \mathcal{X} \to \mathcal{X}$, $p \mapsto p$, est continue.

b) Si deux applications $f : \mathcal{X} \to \mathcal{Y}$ et $g : \mathcal{Y} \to \mathcal{Z}$ sont continues, il en est de même de leur composée

$$g \circ f : \mathcal{X} \to \mathcal{Z}, \quad p \mapsto g(f(p)).$$

Les propriétés a) et b) expriment que l'ensemble TOP de tous les espaces topologiques et de toutes leurs applications continues forme une *catégorie*.

* * *

Si $\mathcal{X}$ et $\mathcal{Y}$ sont des variétés différentiables (de dimensions n et m respectivement), pour tout point $p_0 \in \mathcal{X}$, toute application continue $f : \mathcal{X} \to \mathcal{Y}$ et tout voisinage de coordonnées V du point $f(p)$ dans $\mathcal{Y}$ on peut admettre que le voisinage U du point p_0 dans $\mathcal{X}$ pour lequel $fU \subset V$ est aussi un voisinage de coordonnées. Donc, si

$$h : U \to h(U) \subset \mathbb{R}^n \quad \text{et} \quad k : V \to k(V) \subset \mathbb{R}^m$$

sont des applications de coordonnées, la formule

$$\mathring{f} = k \circ (f \mid_U) \circ h^{-1}$$

définira une application

$$\mathring{f} : h(U) \to k(V) \tag{2}$$

de l'ensemble ouvert $h(U)$ de $\mathbb{R}^n$ dans l'ensemble ouvert $k(V)$ de $\mathbb{R}^m$. Cette application est définie par m fonctions

$$y^j = f^j(x^1, \ldots, x^n), \quad j = 1, \ldots, m, \tag{3}$$

de n variables, exprimant les coordonnées $y^1, \ldots, y^m$ du point $\mathbf{y} = f(\mathbf{x}) \in k(V) \subset \mathbb{R}^m$ au moyen des coordonnées $x^1, \ldots, x^n$ du point $\mathbf{x} \in h(U) \subset \mathbb{R}^n$ (i.e. en d'autres termes, les coordonnées locales du point $q = f(p) \in V$ en fonction de celles du point $p \in U$). On dira que les fonctions (3) *expriment* (ou *définissent*) *l'application f dans les cartes (U, h) et (V, k)* (en coordonnées locales $x^1, \ldots, x^n$ et $y^1, \ldots, y^m$).

Exercice 3. Montrer que si (U', h') et (V', k') sont des cartes telles que $p_0 \in U'$ et $fU' \subset V'$, les fonctions f'^j exprimant l'application f dans les cartes (U', h') et (V', k') sont différentiables au point $h'(p_0)$ si et seulement si les fonctions (3) le sont au point $h(p_0)$.

En ce sens la différentiabilité des fonctions (3) ne dépend pas du choix des cartes (U, h) et (V, k).

Définition 3. On dit qu'une application continue $f: \mathcal{X} \to \mathcal{Y}$ est *différentiable en un point $p_0 \in \mathcal{X}$* si dans certaines (donc dans toutes) cartes (U, h) et (V, k) telles que $p_0 \in U$ et $fU \subset V$, les fonctions (3) définissant l'application f sont des fonctions différentiables (de classe C^r donnée) au point $h(p_0)$.

Une application $f: \mathcal{X} \to \mathcal{Y}$ différentiable en tout point $p \in \mathcal{X}$ est dite *différentiable sur $\mathcal{X}$*.

Exercice 4. Montrer que

a) Pour toute variété différentiable $\mathcal{X}$ l'application identique $\mathrm{id}: \mathcal{X} \to \mathcal{X}$ est différentiable.

b) Si des applications $f: \mathcal{X} \to \mathcal{Y}$ et $g: \mathcal{Y} \to \mathcal{Z}$ sont différentiables, il en est de même de leur composée

$$g \circ f: \mathcal{X} \to \mathcal{Z}, \quad p \mapsto g(f(p)).$$

Par définition cela exprime que *l'ensemble* DIFF *de toutes les variétés différentiables et de toutes leurs applications différentiables forme une catégorie.*

Définition 4. On dit qu'une application $f: \mathcal{X} \to \mathcal{Y}$ de variétés différentiables est un *difféomorphisme* si

1° elle est bijective,

2° différentiable,

3° l'application réciproque $f^{-1}: \mathcal{Y} \to \mathcal{X}$ est différentiable (donc est aussi un difféomorphisme).

Un exemple évident de difféomorphisme est l'application de coordonnées $h: U \to h(U)$. (Question: par quelles fonctions (3) est définie cette application?)

Réciproquement, il est aisé de voir que *pour tout ensemble ouvert U d'une variété différentiable $\mathcal{X}$ et tout difféomorphisme $h: U \to h(U)$,*

$h(U) \subset \mathbb{R}^n$, *le couple* (U, h) *est une carte de* $\mathcal{X}$ (appartient à son atlas maximal). En effet, dire que (U, h) appartient à un atlas maximal revient à dire qu'elle est compatible avec toute carte (V, k) de cet atlas, i.e. les ensembles $h(U \cap V)$ et $k(U \cap V)$ sont ouverts dans $\mathbb{R}^n$ (ou, ce qui est équivalent, dans $h(U)$ et $k(V)$) et l'application $k \circ h^{-1}: h(U \cap V) \to k(U \cap V)$ est un difféomorphisme. Or, la première propriété résulte du fait que l'ensemble $U \cap V$ est ouvert dans V et dans U, et les applications h et k sont des homéomorphismes; la deuxième, du fait que ces deux applications (ou plus exactement leurs restrictions à $U \cap V$) sont des difféomorphismes (comparer avec la démonstration ci-dessus de la proposition 3 et du lemme 1). □

On dit que des variétés $\mathcal{X}$ et $\mathcal{Y}$ sont *difféomorphes* s'il existe au moins un difféomorphisme de $\mathcal{X}$ sur $\mathcal{Y}$. Ces variétés possèdent les mêmes propriétés différentielles (des propriétés exprimées en termes de structure différentiable) et en ce sens sont identiques. En particulier, $\dim \mathcal{X} = \dim \mathcal{Y}$.

Exercice 5. Montrer que les variétés différentiables obtenues en munissant la droite $\mathbb{R}$ de la structure différentiable standard et de la structure de l'exemple 3 de la leçon 6 sont difféomorphes.

Remarque 5. On démontre (faites-le à titre d'exercice) que toute variété différentiable non compacte (resp. compacte) à une dimension vérifiant le deuxième axiome de dénombrabilité est difféomorphe à la droite $\mathbb{R}$ (resp. au cercle S^1) munie de la structure standard. Il est curieux que pour $n = 4$ (et seulement pour $n = 4$) il existe sur $\mathbb{R}^n$ des structures différentiables (qui se construisent très difficilement) compatibles avec la topologie de $\mathbb{R}^n$, pour lesquelles $\mathbb{R}^n$ n'est pas difféomorphe à $\mathbb{R}^n$ muni de la structure standard.

* * *

Soient $\mathcal{X}$ un ensemble, $\mathcal{Y}$ une variété différentiable et $f: \mathcal{X} \to \mathcal{Y}$ une application bijective. *Il existe alors sur* $\mathcal{X}$ *une seule structure différentiable pour laquelle* f *est un difféomorphisme.* Les cartes pour cette structure sont tous les couples $(f^{-1}U, h \circ f)$, où (U, h) est une carte arbitraire de la variété $\mathcal{Y}$. On dit de cette structure qu'elle est *transportée* de $\mathcal{Y}$ sur $\mathcal{X}$ par f.

Il est clair que *les structures différentiables transportées de* $\mathcal{Y}$ *sur* $\mathcal{X}$ *par des applications bijectives* $f: \mathcal{X} \to \mathcal{Y}$ *et* $g: \mathcal{X} \to \mathcal{Y}$ *sont confondues si et seulement si l'application* $g \circ f^{-1}: \mathcal{Y} \to \mathcal{Y}$ *est un difféomorphisme.*

LEÇON 8

Invariance topologique de la dimension des variétés. — Dimension par recouvrements. — Espaces compacts. — Lemme de Lebesgue. — Majoration de la dimension des sous-ensembles compacts de $\mathbb{R}^n$. — Monotonie de la dimension. — Ensembles fermés. — Monotonie de la dimension par ensembles fermés. — Produit direct d'espaces topologiques. — Compacité du produit direct d'espaces compacts.

La conclusion, tirée à la leçon précédente, que la structure d'une variété topologique sur un espace topologique $\mathscr{X}$ n'apporte rien de nouveau à cet espace est en fait un peu hâtive, car à priori il n'est pas exclu qu'un même espace $\mathscr{X}$ puisse posséder des recouvrements ouverts $\{U_\alpha\}$ et $\{V_\beta\}$ tels que tout ensemble U_α soit homéomorphe à un ensemble ouvert $\mathring{U}_\alpha$ de $\mathbb{R}^n$, et tout ensemble V_β, à un ensemble ouvert $\mathring{V}_\beta$ de $\mathbb{R}^m$, $n \neq m$, et par suite, l'espace $\mathscr{X}$ sera muni de deux structures distinctes de variété topologique pour l'une desquelles $\mathscr{X}$ est une variété à n dimensions et pour l'autre une variété à m dimensions. La question de savoir si ceci peut avoir lieu ou non équivaut à la question de savoir *si la dimension est un invariant topologique*, i.e. si des variétés homéomorphes sont nécessairement de même dimension.

Supposons que de tels recouvrements $\{U_\alpha\}$ et $\{V_\beta\}$ existent. Choisissons deux ensembles U_{α_0} et V_{β_0} tels que leur intersection W ne soit pas vide (il est clair que cela est toujours possible) et considérons dans l'espace $\mathbb{R}^n$ l'ensemble O_1, image de W par un homéomorphisme de U_{α_0} sur $\mathring{U}_{\alpha_0}$, et dans l'espace $\mathbb{R}^m$ l'ensemble O_2, image de W par un homéomorphisme de V_{β_0} sur $\mathring{V}_{\beta_0}$. Les ensembles O_1 et O_2 sont ouverts (dans les espaces $\mathbb{R}^n$ et $\mathbb{R}^m$ respectivement) et homéomorphes. Réciproquement, si de tels ensembles existent, ils constituent un exemple de variétés homéomorphes de dimensions différentes.

Donc, la question est de savoir si deux ensembles ouverts appartenant à des espaces $\mathbb{R}^n$ et $\mathbb{R}^m$ de dimensions distinctes ($n \neq m$) peuvent être homéomorphes. On montrera que *la réponse est négative*, et par conséquent, on aura le théorème suivant d'invariance topologique de la dimension des variétés qui dissipe tous les doutes:

Théorème 1. *Les variétés homéomorphes sont de même dimension.*

Pour prouver le théorème 1, on définira pour tout espace topologique $\mathscr{X}$ le nombre entier dim $\mathscr{X}$ appelé *dimension* de $\mathscr{X}$. Par

définition, ce nombre est topologiquement invariant, i.e. le même pour tous espaces homéomorphes. On démontrera en outre que pour tout ensemble $F \subset \mathbb{R}^n$ fermé borné (compact; cf. plus bas) on a l'inégalité

$$\dim F \leqslant n, \tag{1}$$

qui, lorsque F est le cube I^n de $\mathbb{R}^n$, composé des points $\mathbf{t} = (t^1, \ldots, t^n)$ tels que $|t^i| \leqslant 1$ pour tous $i = 1, \ldots, n$, se transforme en l'égalité

$$\dim I^n = n. \tag{2}$$

Ceci suffit déjà pour démontrer le théorème 1.

En effet, soient O_1 et O_2 des ensembles ouverts homéomorphes des espaces $\mathbb{R}^n$ et $\mathbb{R}^m$ respectivement. L'ensemble O_1 étant ouvert, il contient un sous-ensemble fermé et borné F_1 homéomorphe au cube I^n. Soit F_2 l'image de ce sous-ensemble par un homéomorphisme de O_1 sur O_2. Le sous-ensemble F_2 de $\mathbb{R}^m$ est aussi fermé et borné (prouvez-le !). Donc, $\dim F_2 \leqslant m$ en vertu de la formule (1). D'autre part, la fonction dim étant topologiquement invariante, on a $\dim F_2 = \dim F_1 = \dim I^n = n$. Donc, $n \leqslant m$. Les ensembles O_1 et O_2 jouant des rôles symétriques dans ce raisonnement, en les intervertissant on trouve par analogie que $m \leqslant n$. Donc, $n = m$, ce qui prouve le théorème 1. □

Remarque 1. A noter que le nombre $\dim \mathscr{X}$ d'une variété $\mathscr{X}$ (même différentiable) *peut être différent* de sa dimension au sens de la définition 5 de la leçon 6 (cf. exemple 4 plus bas). On désignera provisoirement (dans cette leçon seulement) par $\dim' \mathscr{X}$ la dimension de la variété $\mathscr{X}$ au sens de la définition 5 de la leçon 6. Plus bas (cf. remarque 3) on caractérisera topologiquement ce nombre (et on démontrera par là même le théorème 1 une fois de plus).

* * *

Introduisons la fonction dim à l'aide des définitions suivantes:

Définition 1. On dit qu'un recouvrement $\{U_\alpha\}$ d'un espace $\mathscr{X}$ est *inscrit* dans un recouvrement $\{V_\beta\}$ si pour tout α il existe un β tel que $U_\alpha \subset V_\beta$.

Définition 2. On dit qu'un recouvrement $\{U_\alpha\}$ d'un espace $\mathscr{X}$ est de *multiplicité* $\leqslant n + 1$ si l'intersection de toute famille à $n + 2$ ensembles de $\{U_\alpha\}$ est vide. Si le recouvrement $\{U_\alpha\}$ est de multiplicité $\leqslant n + 1$ mais pas de multiplicité $\leqslant n$ (i.e. il contient une famille à $n + 1$ ensembles d'intersection non vide), on dit qu'il est *de multiplicité* $n + 1$.

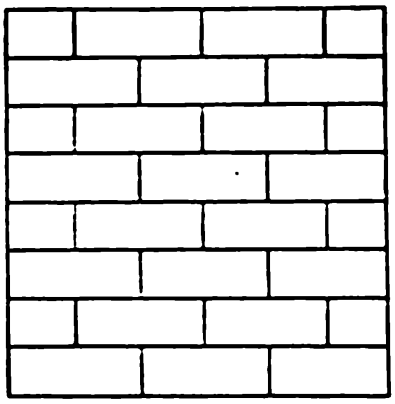
Définition 3. On dit que $\dim \mathscr{X} \leqslant n$ si dans tout recouvrement ouvert fini d'un espace $\mathscr{X}$ on peut inscrire un recouvrement ouvert fini de multiplicité $\leqslant n + 1$. Si $\dim \mathscr{X} \leqslant n$ mais il est faux que $\dim \mathscr{X} \leqslant n + 1$, on dit que $\dim \mathscr{X} = n$.

Exemple 1. Un espace discret est de toute évidence de dimension 0.

Exemple 2. L'ensemble de tous les nombres rationnels (pour la topologie induite par celle de la droite réelle) est aussi de dimension 0 (dans tout recouvrement ouvert fini de cet ensemble on peut inscrire un recouvrement fini composé d'intervalles disjoints, donc de multiplicité 1).

Exemple 3. L'ensemble des nombres irrationnels est aussi de dimension 0 pour des raisons analogues.

Exemple 4. Soit $\mathscr{X}$ une droite non séparée avec une singularité de multiplicité $n+1$ en 0 et soit $\{U_\alpha\}$ un recouvrement ouvert fini de l'espace $\mathscr{X}$ tel que les $n+1$ points en lesquels se désagrège le



point 0 appartiennent à des éléments distincts de ce recouvrement. La multiplicité de ce recouvrement et aussi de tout recouvrement ouvert inscrit dans $\{U_\alpha\}$ sera $\geqslant n+1$. Ceci montre que $\dim \mathscr{X} \geqslant n$ (alors que $\dim' \mathscr{X} = 1$).

Donc, dans le cas des variétés non séparées, l'invariant dim ne reflète pas adéquatement l'idée intuitive de la dimension.

Remarque 2. On ne sait pas si l'égalité $\dim \mathscr{X} = \dim' \mathscr{X}$ a lieu pour toute variété séparée. On démontre que $\dim \mathscr{X} = \dim' \mathscr{X}$ si la variété $\mathscr{X}$ est paracompacte (les variétés paracompactes seront définies à la leçon 23).

Exemple 5. Il est évident que dans tout recouvrement ouvert fini du segment $I = [-1, 1]$ on peut inscrire un recouvrement constitué d'intervalles empiétant légèrement les uns sur les autres (donc de multiplicité 2). Par conséquent, $\dim I \leqslant 1$. Comme l'égalité $\dim I = 0$ est visiblement impossible (le segment I ne peut être recouvert par une famille finie d'intervalles disjoints de longueur < 1), on en déduit que $\dim I = 1$ (cf. formule (2)).

Exercice 1. Donner une forme rigoureuse au raisonnement produit.

Exemple 6. En recouvrant le carré $I^2 = \{(x, y) \in \mathbb{R}^2, -1 \leqslant x \leqslant 1, -1 \leqslant y \leqslant 1\}$ par des pavés en quiconque et en augmentant

légèrement les dimensions de chaque pavé, on obtient un recouvrement ouvert fini de multiplicité 3. On démontre (cf. plus bas, mais nous recommandons au lecteur de le faire immédiatement) qu'un tel pavage peut être inscrit dans tout recouvrement ouvert fini du carré. Donc, $\dim I^2 \leqslant 2$. (Nous verrons que la démonstration de l'inégalité inverse $\dim I^2 \geqslant 2$ est bien plus compliquée.)

* * *

On se propose de généraliser la construction de l'exemple 6 à un cube à n dimensions

$$I^n = \{(t_1, \ldots, t_n) \in \mathbb{R}^n, \ -1 \leqslant t_1 \leqslant 1, \ldots, -1 \leqslant t_n \leqslant 1\}.$$

Pour cela il nous faut effectuer un certain travail préparatoire.

Définition 4. On dit qu'un espace topologique $\mathscr{X}$ est *compact* (d'après Bourbaki, *quasi compact*) si de tout recouvrement ouvert $\{U_\alpha\}$ de $\mathscr{X}$ on peut extraire un sous-recouvrement fini.

Il est aisé de voir que *tout sous-ensemble infini d'un espace compact admet un point d'accumulation* (i.e. un point dont chaque voisinage contient une infinité de points de ce sous-ensemble). En effet, dans le cas contraire, chaque point de l'espace posséderait un voisinage contenant un nombre fini de points du sous-ensemble. Ces voisinages formeraient un recouvrement ouvert dont aucune sous-famille finie ne serait un recouvrement (puisque le nombre total de points de ce sous-ensemble contenu dans les éléments de cette sous-famille est fini). Vu qu'un tel recouvrement ne peut exister dans un espace compact, on en déduit que les points d'accumulation existent. □

Un sous-ensemble A d'un espace topologique est dit *compact* s'il l'est pour la topologie induite (en tant que sous-espace).

On appelle *recouvrement ouvert* d'un sous-espace A de $\mathscr{X}$ une famille $\{U_\alpha\}$ d'ensembles ouverts de $\mathscr{X}$ telle que

$$A \subset \bigcup_\alpha U_\alpha.$$

Dans ce cas, les intersections $A \cap U_\alpha$ forment un recouvrement ouvert de A en tant qu'espace topologique (pour la topologie induite), et tout recouvrement de A peut être ainsi obtenu (mais de façon non unique). Donc, *un sous-espace est compact si et seulement si de tout recouvrement ouvert on peut extraire un sous-recouvrement fini.*

Le théorème classique de Heine-Borel nous dit que *tout sous-espace fermé borné de* $\mathbb{R}^n$ *est compact.* En particulier, *le cube* I^n *est compact.*

* * *

On rappelle que le *diamètre* d'un sous-ensemble K d'un espace métrique $\mathscr{X}$ est par définition le nombre

$$d(K) = \sup_{p,\, q \in K} \rho(p, q),$$

où ρ est une métrique sur $\mathscr{X}$.

Lemme 1 (lemme de Lebesgue). *Pour tout recouvrement ouvert $\{U_\alpha\}$ d'un espace métrique compact $\mathscr{X}$ il existe un nombre $\varepsilon > 0$ tel que tout sous-ensemble K de $\mathscr{X}$ dont le diamètre est $< \varepsilon$ est contenu dans un élément du recouvrement $\{U_\alpha\}$.*

Démonstration. Si un tel ε n'existe pas, pour tout $n > 0$ on peut exhiber un sous-ensemble K_n de $\mathscr{X}$, de diamètre $< 1/n$, n'appartenant à aucun U_α. Choisissons un point arbitraire p_n de K_n et considérons l'ensemble $\{p_n\}$. L'espace $\mathscr{X}$ étant compact, cet ensemble possède au moins un point d'accumulation p_0. Soit $p_0 \in U_{\alpha_0}$ et soit d la distance de p_0 à $\mathscr{X} \setminus U_{\alpha_0}$ (i.e. $d = \inf \rho(p_0, p)$, où $p \in \mathscr{X} \setminus U_\alpha$). Si $n > 2/d$ et $\rho(p_0, p_n) < d/2$, alors

$$\rho(p_0, p) \leqslant \rho(p_0, p_n) + \rho(p_n, p) \leqslant \frac{d}{2} + \frac{1}{n} < d$$

pour tout point $p \in K_n$, et donc, contrairement à l'hypothèse, $K_n \subset U_{\alpha_0}$. □

Le nombre $\varepsilon > 0$ du lemme s'appelle *nombre de Lebesgue* du recouvrement $\{U_\alpha\}$.

Un recouvrement $\{V_\beta\}$ d'un espace métrique $\mathscr{X}$ s'appelle ε-recouvrement si $d(V_\beta) < \varepsilon$ pour tout β.

Corollaire 1. *Si ε est le nombre de Lebesgue d'un recouvrement $\{U_\alpha\}$, tout ε-recouvrement ouvert fini $\{V_\beta\}$ d'un espace métrique compact $\mathscr{X}$ est inscrit dans $\{U_\alpha\}$.* □

Corollaire 2. *L'inégalité* $\dim \mathscr{X} \leqslant n$ *est satisfaite pour un espace métrique compact $\mathscr{X}$ si et seulement si pour tout $\varepsilon > 0$ il existe un ε-recouvrement ouvert fini de $\mathscr{X}$ de multiplicité $\leqslant n + 1$.* □

* * *

Désormais, nous sommes en mesure de généraliser l'exemple 6 à tout n (et à tout ensemble fermé borné $F \subset \mathbb{R}^n$).

Proposition 1. *Pour tout ensemble fermé borné $F \subset \mathbb{R}^n$ on a l'inégalité*

$$\dim F \leqslant n.$$

Démonstration. Soit $N > 0$. Par récurrence sur n construisons pour tout $n \geqslant 1$ un recouvrement spécial de $\mathbb{R}^n$ composé d'ensembles fermés, que l'on appellera *pavage de rang N*. Ceci étant,

les points de l'espace $\mathbb{R}^n$, $n \geqslant 2$, seront identifiés aux couples $(\mathbf{t}, t)$, où $\mathbf{t} \in \mathbb{R}^{n-1}$ et $t \in \mathbb{R}$.

Par définition, le pavage de rang N de la droite $\mathbb{R}$ est composé des segments $[(a-1)/N, a/N]$, $-\infty < a < +\infty$.

Supposons qu'on ait déjà construit un pavage de rang N de l'espace $\mathbb{R}^{n-1}$, $n \geqslant 2$. Pour pavage de rang N de l'espace $\mathbb{R}^n$ prenons le recouvrement composé de tous les ensembles de la forme

$$K \times [(a-1)/N, a/N] = \{(\mathbf{t}, t) \in \mathbb{R}^n, \mathbf{t} \in K, (a-1)/N \leqslant t \leqslant a/N\},$$

où a est un entier arbitraire, et K un élément du pavage de rang N de l'espace $\mathbb{R}^{n-1}$ si a est impair et un élément du pavage de rang N de l'espace $\mathbb{R}^{n-1}$ soumis à une $(1/2N, \ldots, 1/2N)$-translation si a est pair.

Par récurrence sur n on démontre immédiatement que:

a) un pavage de rang N de l'espace $\mathbb{R}^n$ est un recouvrement de l'espace $\mathbb{R}^n$ formé de cubes d'arête $1/N$;

b) tout point de l'espace $\mathbb{R}^n$ appartient au plus à $n+1$ cubes d'un pavage de rang N, et de plus un point appartient à $n+1$ cubes si et seulement s'il est de la forme

$$\left(\frac{a_1}{2N}, \ldots, \frac{a_{n-1}}{2N}, \frac{a_n}{N}\right),$$

où $a_1, \ldots, a_{n-1}, a_n$ sont des entiers.

Le diamètre d'un cube d'arête a de $\mathbb{R}^n$ étant visiblement égal à $a\sqrt{n}$, on voit que le pavage de rang N est un $(\sqrt{n}/N)$-recouvrement *fermé* (i.e. est composé d'ensembles fermés) de $\mathbb{R}^n$ de multiplicité $n+1$. Pour obtenir un recouvrement ouvert, choisissons un nombre $\delta > 0$ et remplaçons chaque cube du pavage par son δ-voisinage (dont le diamètre est de toute évidence égal à $2\delta + (\sqrt{n}/N)$). Pour δ assez petit (plus exactement pour $\delta < 1/2N$) la multiplicité du recouvrement ne changera pas. Puisque pour tout $\varepsilon > 0$ il existe des N et δ tels que $\varepsilon > 2\delta + (\sqrt{n}/N)$, on voit donc que *pour tout* $\varepsilon > 0$ *il existe un* ε*-recouvrement ouvert de l'espace* $\mathbb{R}^n$ *de multiplicité* $n+1$.

Tout sous-ensemble fermé borné (i.e. compact) F de l'espace $\mathbb{R}^n$ n'est intersecté que par un nombre fini d'éléments de ce recouvrement, donc ce recouvrement découpe sur F un ε-recouvrement ouvert fini de multiplicité $\leqslant n+1$. Par conséquent, $\dim F \leqslant n$. □

Des formules (1) et (2) indispensables à la démonstration du théorème 1, nous avons déjà établi la première. Pour achever la démonstration du théorème 1, il ne nous reste donc qu'à prouver l'inégalité $\dim I^n \geqslant n$ (ce qui en fait est la partie la plus difficile de la démonstration). Pour ne pas suspendre l'exposé, nous le ferons à la prochaine leçon et, pour l'instant, discutons plus en détail la formule (1).

* * *

Par une meilleure exploitation de la construction des pavages, on démontre que la formule (1) est valable pour tous sous-ensembles F de $\mathbb{R}^n$ (pas nécessairement fermés et bornés).

Exercice 2. Prouver la formule (1) pour tout sous-ensemble $F \subset \mathbb{R}^n$.

Pour $F = \mathbb{R}^n$ en particulier, on trouve que

$$\dim \mathbb{R}^n \leqslant n. \tag{3}$$

Donc, la formule (1) est la conséquence de l'inégalité formellement plus exacte

$$\dim F \leqslant \dim \mathbb{R}^n \tag{4}$$

(en effet, on verra plus bas que dans la formule (3) on a l'égalité

$$\dim \mathbb{R}^n = n, \tag{5}$$

donc les inégalités (1) et (4) sont équivalentes).

La formule (4) affirme que la dimension de tout sous-espace F de l'espace topologique $\mathbb{R}^n$ est inférieure à la dimension de cet espace, ce qui est bien en accord avec notre intuition. Il est naturel d'espérer que la propriété de monotonie de la dimension

$$\dim F \leqslant \dim \mathscr{X} \tag{6}$$

ait lieu pour un sous-espace F de tout espace topologique $\mathscr{X}$. Mais nous verrons sur un exemple de la leçon 10 que *l'inégalité* (6) *est généralement fausse* et pour qu'elle soit vraie il faut imposer à F ou à $\mathscr{X}$ des conditions subsidiaires.

* * *

Définition 5. On dit qu'un sous-ensemble F d'un espace topologique $\mathscr{X}$ est *fermé* si son complémentaire est ouvert.

Pour les sous-ensembles des espaces métriques (et en particulier pour les sous-ensembles de l'espace $\mathbb{R}^n$) cette définition est connue du cours d'analyse (et nous l'avons déjà utilisée pour ces espaces).

Exercice 3. Montrer que *tout singleton d'une variété $\mathscr{X}$ est fermé* (i.e. tout point de $\mathscr{X}$ est fermé).

Exercice 4. Montrer que *tout sous-ensemble fermé d'un espace compact est compact.* [N o t a. Se servir du fait qu'en ajoutant à un recouvrement ouvert d'un ensemble F l'ensemble ouvert $\mathscr{X} \setminus F$, on obtient un recouvrement ouvert de l'espace $\mathscr{X}$ tout entier.]

Il est clair que:

1° L'ensemble vide $\varnothing$ et l'espace $\mathscr{X}$ sont fermés.

2° L'intersection de toute famille d'ensembles fermés est fermée.

3° La réunion de toute famille finie d'ensembles fermés est fermée.

Comparer avec les propriétés homologues 1° à 3° des ensembles ouverts à la leçon 7.

La propriété 2° des ensembles ouverts entraîne immédiatement que *pour tout sous-ensemble A d'un espace topologique, il existe un plus petit ensemble fermé contenant A* (ce sera l'intersection de tous les ensembles fermés contenant A). Cet ensemble fermé s'appelle *adhérence* de l'ensemble A et se note $\overline{A}$.

Il est clair qu'*un point $p \in \mathscr{X}$ appartient à $\overline{A}$ si et seulement si tout voisinage de p rencontre A*. Ceci exprime que

$$\mathscr{X} \setminus \overline{A} = \operatorname{Int}(\mathscr{X} \setminus A).$$

[On remarquera que $\operatorname{Int} A$ est justiciable de la formule duale $\mathscr{X} \setminus \operatorname{Int} A = \overline{\mathscr{X} \setminus A}$.]

Exercice 5. Montrer que les propriétés suivantes de l'application $f : \mathscr{X} \to \mathscr{Y}$ sont équivalentes:

a) L'application f est continue.

b) Pour tout ensemble fermé $C \subset \mathscr{Y}$ l'ensemble $f^{-1} C \subset \mathscr{X}$ est fermé.

c) Pour tout ensemble $A \subset \mathscr{X}$ on a l'inclusion

$$f(\overline{A}) \subset \overline{f(A)}.$$

d) Pour tout ensemble $B \subset \mathscr{Y}$ on a l'inclusion

$$\overline{f^{-1}(B)} \subset f^{-1}(\overline{B}).$$

Remarque 3. Nous avons vu que l'invariant dim ne correspondait pas toujours à notre intuition de la dimension. On pourrait tenter d'arranger la situation en introduisant un nouvel invariant dim'. Pour tout espace topologique $\mathscr{X}$ on admettra que $\dim' \mathscr{X} \leqslant n$ si pour tout voisinage U d'un point $p \in \mathscr{X}$ il existe un voisinage V de p contenu dans U, tel que pour son adhérence $\overline{V}$ dans U l'on ait l'inégalité $\dim \overline{V} \leqslant n$. (La mention « dans U » est essentielle ici, car l'adhérence du voisinage V dans $\mathscr{X}$ peut être bien plus grande.) Si $\dim' \mathscr{X} \leqslant n$ mais $\dim' \mathscr{X} \leqslant n - 1$ est faux, alors par définition $\dim' \mathscr{X} = n$. Puisque tout point d'un ensemble ouvert $U \subset \mathbb{R}^n$ possède un voisinage dont l'adhérence est homéomorphe au cube I^n (on se sert de la formule (2) qui reste à prouver), la dimension $\dim' \mathscr{X}$ de toute variété $\mathscr{X}$ sera confondue avec celle de la définition 5 de la leçon 6 (cf. remarque 1 plus haut). Les conditions nécessaires et suffisantes pour que l'égalité

$$\dim \mathscr{X} = \dim' \mathscr{X}$$

ait lieu sont encore inconnues (cf. plus haut exemple 4 et remarque 2).

* * *

Comme déjà dit, la dimension d'un sous-espace peut être supérieure à celle de l'espace tout entier.

Mais ceci est impossible pour les sous-espaces fermés.

Proposition 2. *Pour tout sous-espace fermé F d'un espace topologique $\mathcal{X}$.*

$$\dim F \leqslant \dim \mathcal{X}. \tag{7}$$

Démonstration. Soit $\{U_\alpha\}$ un recouvrement ouvert fini d'un sous-espace F. Dire que U_α est ouvert dans F revient à dire qu'il existe dans $\mathcal{X}$ des ensembles ouverts U'_α tels que $U_\alpha = U'_\alpha \cap F$ pour tout α. Les ensembles U'_α forment avec l'ensemble $\mathcal{X} \setminus F$ un recouvrement ouvert fini de l'espace $\mathcal{X}$. Il existe donc un recouvrement ouvert fini $\{V_\beta\}$ de multiplicité $\leqslant n+1$, où $n = \dim \mathcal{X}$, inscrit dans ce recouvrement. Considérons tous les ensembles non vides de la forme $V'_\beta \cap F$. Il est clair que ces ensembles forment un recouvrement ouvert fini d'un sous-espace F de multiplicité $\leqslant n+1$, inscrit dans le recouvrement $\{U_\alpha\}$. Donc, dans tout recouvrement ouvert fini de F on peut inscrire un recouvrement ouvert fini de multiplicité $\leqslant n+1$. Donc, $\dim F \leqslant n$. □

Dans l'exemple qui sera construit à la leçon 10, le sous-espace F sera ouvert (et l'espace $\mathcal{X}$ séparé et compact).

On démontre (c'est un théorème difficile) que, pour que l'inégalité (6) ait lieu pour tout sous-espace F, il suffit que l'espace $\mathcal{X}$ soit métrisable et satisfasse au deuxième axiome de dénombrabilité (d'une façon générale, les espaces métrisables satisfaisant au deuxième axiome de dénombrabilité sont dans un certain sens les domaines les plus naturels de construction de la théorie de la dimension; en dehors de cette classe d'espaces, presque aucune propriété « naturelle » de la dimension n'est généralement vraie). Donc, dans l'exemple de la leçon 10, à priori l'espace $\mathcal{X}$ soit n'est pas métrisable, soit ne satisfait pas au deuxième axiome de dénombrabilité. (En effet, on démontre sans peine qu'il n'est pas métrisable et ne satisfait pas au deuxième axiome de dénombrabilité. Ceci ne l'empêche pas d'être séparé et compact.)

Remarque 4. De la proposition 2 il résulte en particulier que $\dim I^n \leqslant \dim \mathbb{R}^n$. D'après la formule (2) qui sera prouvée à la leçon 9, ceci exprime que $n \leqslant \dim \mathbb{R}^n$. Combiné à la formule (3) qui n'a pas encore été prouvée, ceci nous donne l'égalité (5).

* * *

Outre l'approche de la démonstration de la formule (3) esquissée plus haut, il en existe une autre plus conceptuelle qui est basée sur une construction topologique simple mais importante exposée pour des cas particuliers en analyse.

Soient $\mathscr{X}$ et $\mathscr{Y}$ des espaces topologiques et soit $\mathscr{X} \times \mathscr{Y}$ l'ensemble de tous les couples (p, q), où $p \in \mathscr{X}$ et $q \in \mathscr{Y}$. Si $U \subset \mathscr{X}$ et $V \subset \mathscr{Y}$, alors $U \times V \subset \mathscr{X} \times \mathscr{Y}$ et il est clair que l'ensemble de tous les sous-ensembles de la forme $U \times V$, où U est ouvert dans $\mathscr{X}$ et V ouvert dans $\mathscr{Y}$, est stable pour les intersections. Cet ensemble est donc une base pour une certaine topologie de $\mathscr{X} \times \mathscr{Y}$.

Définition 6. L'espace topologique obtenu s'appelle *produit direct* (ou *cartésien*) des espaces $\mathscr{X}$ et $\mathscr{Y}$ et se note par le même symbole $\mathscr{X} \times \mathscr{Y}$.

La topologie de l'espace $\mathscr{X} \times \mathscr{Y}$ s'appelle *produit direct* des topologies des espaces $\mathscr{X}$ et $\mathscr{Y}$.

Il est aisé de voir que *le produit direct $\mathscr{X} \times \mathscr{Y}$ d'espaces séparés $\mathscr{X}$ et $\mathscr{Y}$ est séparé*. En effet, soient (p_1, q_1) et (p_2, q_2) deux points distincts de l'espace $\mathscr{X} \times \mathscr{Y}$. Il s'ensuit que soit $p_1 \neq p_2$, soit $q_1 \neq q_2$. Admettons pour plus de clarté que $p_1 \neq p_2$. L'espace $\mathscr{X}$ étant par hypothèse séparé, les points p_1 et p_2 possèdent dans $\mathscr{X}$ des voisinages U_1 et U_2 disjoints. Les ensembles $U_1 \times \mathscr{Y}$ et $U_2 \times \mathscr{Y}$ seront alors des voisinages disjoints des points (p_1, q_1) et (p_2, q_2) dans le produit $\mathscr{X} \times \mathscr{Y}$. □

On définit par analogie le produit direct $\mathscr{X}_1 \times \ldots \times \mathscr{X}_n$ de toute famille f i n i e d'espaces topologiques. Il est aussi séparé si les espaces $\mathscr{X}_1, \ldots, \mathscr{X}_n$ le sont.

Exercice 6. Montrer que

$$\mathbb{R}^n = \mathbb{R} \times \ldots \times \mathbb{R} \ (n \text{ fois}).$$

[N o t a. Pour toute boule de l'espace $\mathbb{R}$ il existe un cube inscrit et un cube circonscrit de faces parallèles aux plans de coordonnées.]

Pour que des ensembles $U \times V$ forment une base de l'espace $\mathscr{X} \times \mathscr{Y}$, il n'est pas nécessaire que U et V parcourent tous les ensembles ouverts des espaces $\mathscr{X}$ et $\mathscr{Y}$ respectivement. Il suffit de toute évidence que les ensembles U parcourent une base de $\mathscr{X}$, et les ensembles V, une base de $\mathscr{Y}$. Donc, *si des espaces satisfont au deuxième axiome de dénombrabilité, il en est de même de leur produit direct.*

En particulier, on voit encore que l'espace $\mathbb{R}^n$ satisfait au deuxième axiome de dénombrabilité.

L'examen d'exemples élémentaires (disons des cubes I^n) nous incite à penser que pour deux espaces quelconques $\mathscr{X}$ et $\mathscr{Y}$ on doit avoir

$$\dim (\mathscr{X} \times \mathscr{Y}) = \dim \mathscr{X} + \dim \mathscr{Y}. \tag{8}$$

Mais des exemples que nous ne pouvons développer en raison de leur complexité montrent que *l'égalité* (8) *n'est même pas vérifiée pour des espaces métriques compacts.*

Exercice 7. Montrer que tout espace métrique compact satisfait au deuxième axiome de dénombrabilité.

Bien plus, même l'inégalité

$$\dim (\mathscr{X} \times \mathscr{Y}) \leqslant \dim \mathscr{X} + \dim \mathscr{Y} \tag{9}$$

peut ne pas avoir lieu pour des espaces topologiques $\mathscr{X}$ et $\mathscr{Y}$ quelconques. Cependant, on démontre que *l'inégalité* (9) *est satisfaite dans la classe des espaces métriques vérifiant le deuxième axiome de dénombrabilité* (ainsi que dans la classe des espaces compacts séparés). Puisque, comme il est aisé de le voir, $\dim \mathbb{R} = 1$ (cf. exemple 5 ci-dessus), ceci prouve en particulier l'inégalité (3). Mais la démonstration de la formule (9) pour les espaces métriques vérifiant le deuxième axiome de dénombrabilité est assez compliquée et nous glisserons dessus faute de place.

* * *

La véracité de la formule (9) dans la classe des espaces compacts séparés passe en particulier par celle de la proposition suivante:

Proposition 3. *Le produit direct d'espaces topologiques compacts est compact.*

Avant d'aborder la démonstration faisons quelques remarques sur la projection

$$\mathscr{X} \times \mathscr{Y} \to \mathscr{X}, \quad (p, q) \mapsto p. \tag{10}$$

L'image réciproque par la projection (10) d'un ensemble ouvert quelconque $U \subset \mathscr{X}$ est l'ensemble $U \times \mathscr{Y}$ (ouvert par définition). Donc, *la projection* (10) *est une application continue.*

Bien plus, étant donné que tout ensemble de la forme $U \times V$ se projette sur U, *l'image de tout ensemble ouvert de l'espace* $\mathscr{X} \times \mathscr{Y}$ *par la projection* (10) *est un ensemble ouvert de l'espace* $\mathscr{X}$ (i.e., la projection (10) est une application ouverte).

Il est évident que des assertions analogues sont valables pour la projection $\mathscr{X} \times \mathscr{Y} \to \mathscr{Y}$, $(p, q) \mapsto q$, ainsi que pour les projections $\mathscr{X}_1 \times \ldots \times \mathscr{X}_n \to \mathscr{X}_k$ du produit d'un nombre quelconque d'espaces sur l'un des facteurs de ce produit.

Dans le même temps, la projection (10) peut généralement appliquer un ensemble fermé dans un ensemble non fermé (par exemple, la projection $\mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x$, applique l'hyperbole $xy = 1$, qui est un ensemble fermé du plan $\mathbb{R}^2$, dans l'ensemble non fermé $\mathbb{R} \setminus \{0\}$ de l'axe $\mathbb{R}$). Cependant, *si l'espace* $\mathscr{Y}$ *est compact, l'image de tout ensemble fermé de l'espace* $\mathscr{X}$ *par la projection* (10) *est un ensemble fermé de* $\mathscr{X}$ (autrement dit, la projection (10) est une *application fermée*). En effet, supposons qu'un point $p_0 \in \mathscr{X}$ n'appartient pas à la projection $\operatorname{pr} F \subset \mathscr{X}$ d'un ensemble fermé $F \subset \mathscr{X} \times \mathscr{Y}$, i.e. que $\{p_0\} \times \mathscr{Y} \subset (\mathscr{X} \times \mathscr{Y}) \setminus F$. L'ensemble $(\mathscr{X} \times \mathscr{Y}) \setminus F$ étant ouvert, donc composé uniquement de points intérieurs, on en déduit que pour tout point $q \in \mathscr{Y}$ il existe un voisinage $V_q \subset \mathscr{Y}$ et un voisinage

$U(q) \subset \mathscr{X}$ de p_0 tels que $U(q) \times V_q \subset (\mathscr{X} \times \mathscr{Y}) \setminus F$. Les voisinages V_q, $q \in \mathscr{Y}$, forment tous un recouvrement ouvert de l'espace $\mathscr{Y}$, donc — puisque $\mathscr{Y}$ est compact — de ce recouvrement on peut extraire un sous-recouvrement fini $V_{q_1}, \ldots, V_{q_n}$. Soit

$$U = U(q_1) \cap \ldots \cap U(q_n).$$

L'ensemble U est ouvert, contient le point p_0 et possède la propriété suivante: aucun point du produit $\mathscr{X} \times \mathscr{Y}$ de la forme (p, q), où $p \in U$, $q \in \mathscr{Y}$, n'appartient à F (si $q \in V_{q_i}$, alors $(p, q) \in U(q_i) \times \times V_{q_i} \subset (\mathscr{X} \times \mathscr{Y}) \setminus F$). Ceci exprime que $\{p\} \times \mathscr{Y} \subset (\mathscr{X} \times \mathscr{Y}) \setminus F$, i.e. que $p \notin \operatorname{pr} F$. Nous avons ainsi prouvé que tout point $p_0 \notin \operatorname{pr} F$ possède un voisinage $U \subset \mathscr{X}$ tel que $U \subset \mathscr{X} \setminus \operatorname{pr} F$, i.e. le point p_0 est un point intérieur de l'ensemble $\mathscr{X} \setminus \operatorname{pr} F$. Donc, cet ensemble est ouvert et, par suite, l'ensemble $\operatorname{pr} F$ est fermé. □

Nous sommes maintenant en mesure de prouver la proposition 3.

Démonstration de la proposition 3. Supposons que les espaces $\mathscr{X}$ et $\mathscr{Y}$ sont compacts et soit $\{W_\alpha\}$ un recouvrement ouvert de l'espace $\mathscr{X} \times \mathscr{Y}$. On dira (uniquement dans cette démonstration !) qu'un ensemble ouvert $O \subset \mathscr{X}$ est *distingué* si le sous-ensemble $O \times \mathscr{Y}$ du produit $\mathscr{X} \times \mathscr{Y}$ est contenu dans la réunion d'une sous-famille finie du recouvrement $\{W_\alpha\}$. Pour tout point $p \in \mathscr{X}$ le sous-espace $\{p\} \times \mathscr{Y}$ du produit $\mathscr{X} \times \mathscr{Y}$ est homéomorphe (prouvez-le) à l'espace $\mathscr{Y}$, donc est compact. Par conséquent, on peut le recouvrir par une sous-famille finie du recouvrement $\{W_\alpha\}$. Soit F le sous-ensemble fermé de l'espace $\mathscr{X} \times \mathscr{Y}$, complémentaire, relativement à $\mathscr{X} \times \mathscr{Y}$, de la réunion G de tous les éléments de cette sous-famille. L'espace $\mathscr{Y}$ étant compact par hypothèse, la projection $\operatorname{pr} F$ est fermée dans $\mathscr{X}$ d'après ce qui a été prouvé, donc son complémentaire O est ouvert. D'autre part, l'ensemble $O \times \mathscr{Y}$ est contenu dans G par construction. Donc, l'ensemble O est distingué. Puisque $p \in O$, ceci prouve que tous les ensembles distingués forment un recouvrement de l'espace $\mathscr{X}$. Cet espace étant compact, on peut extraire de ce recouvrement un sous-recouvrement fini, i.e. trouver une famille finie d'ensembles distingués recouvrant $\mathscr{X}$. Or, il est clair que la réunion de toute famille finie d'ensembles distingués est un ensemble distingué. Donc, $\mathscr{X}$ est distingué et, par suite, le recouvrement $\{W_\alpha\}$ contient un sous-recouvrement fini. □

La démonstration de l'inégalité (9) pour les espaces compacts séparés est aussi trop compliquée pour être exposée ici.

LEÇON 9

Théorème du tambour. — Théorème de Brouwer du point fixe. — Théorème des cloisons dans le cube. — Espaces normaux et complètement normaux. — Prolongement des cloisons. — Théorème de Lebesgue de recouvrements du cube. — Minoration de la dimension du cube.

A la leçon précédente nous avons montré que dim $I^n \leqslant n$. Le principal objectif de cette leçon est de prouver l'inégalité inverse et de montrer ainsi que

$$\dim I^n = n \quad \text{pour tout } n \geqslant 0.$$

Pour cela il nous faut commencer de loin.

Définition 1. Un sous-espace A d'un espace topologique $\mathscr{X}$ s'appelle *retract* de $\mathscr{X}$ s'il existe une application continue

$$r : \mathscr{X} \to A$$

(appelée *application rétractante* ou tout simplement *rétraction*) telle que $r(a) = a$ pour tout $a \in A$.

Rappelons que par $\mathbb{B}^n$ on désigne la *boule unité* de l'espace $\mathbb{R}^n$, i.e. l'ensemble des points $\mathbf{x} \in \mathbb{R}^n$ tels que $|\mathbf{x}| \leqslant 1$ (où comme toujours $|\mathbf{x}| = \sqrt{x_1^2 + \ldots + x_n^2}$ si $\mathbf{x} = (x_1, \ldots, x_n)$), et par $\mathbb{S}^{n-1}$ la sphère unité, i.e. le sous-ensemble de $\mathbb{B}^n$ composé des points $\mathbf{x} \in \mathbb{R}^n$ tels que $|\mathbf{x}| = 1$.

Théorème 1. *La sphère $\mathbb{S}^{n-1}$ n'est pas un retract de la boule $\mathbb{B}^n$.*

Pour $n = 2$, ce théorème explique théoriquement qu'il est possible de tendre une membrane sur un cercle, i.e. de faire un tambour. Pour cette raison le théorème 1 est appelé généralement théorème du tambour.

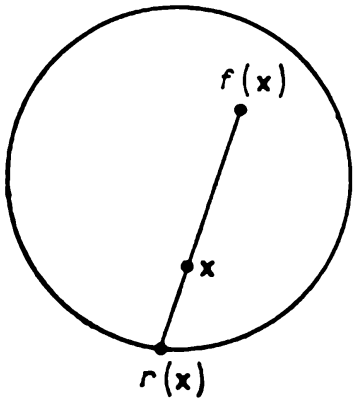
Contre toute attente, la démonstration de ce théorème évident est assez compliquée et implique l'utilisation d'une foule d'idées nouvelles. Nous l'ajournerons donc à la leçon 24.

* * *

Une conséquence immédiate du théorème 1 est le théorème de Brouwer du point fixe suivant:

Théorème 2. *Pour toute application continue $f : \mathbb{B}^n \to \mathbb{B}^n$ il existe un point $\mathbf{x} \in \mathbb{B}^n$ tel que $f(\mathbf{x}) = \mathbf{x}$.*

D é m o n s t r a t i o n. Si $f(\mathbf{x}) \neq \mathbf{x}$, il existe une droite passant par les points $\mathbf{x}$ et $f(\mathbf{x})$. Soit $r(\mathbf{x})$ celui des deux points d'intersection de cette droite avec la sphère $\mathbb{S}^{n-1}$, qui n'est pas séparé du point $\mathbf{x}$ par le point $f(\mathbf{x})$. Si $f(\mathbf{x}) \neq \mathbf{x}$ pour tous les points $\mathbf{x}$, cette construction définit une application $r: \mathbb{B}^n \to \mathbb{S}^{n-1}$ qui est visiblement continue. Si $\mathbf{x} \in \mathbb{S}^{n-1}$, le point d'intersection non séparé du point $\mathbf{x}$ est le point $\mathbf{x}$ lui-même. Donc, $r(\mathbf{x}) = \mathbf{x}$, i.e. r est une rétrac-



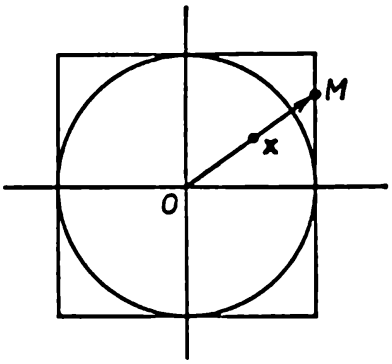
tion. Puisque l'existence d'une telle rétraction contredit le théorème 1, l'inégalité $f(\mathbf{x}) \neq \mathbf{x}$ ne peut avoir lieu pour tous les points $\mathbf{x} \in \mathbb{B}^n$. □

Corollaire 1. *Pour toute application continue* $f: I^n \to I^n$ *il existe un point* $\mathbf{t}_0 \in I^n$ *tel que* $f(\mathbf{t}_0) = \mathbf{t}_0$.

D é m o n s t r a t i o n. Il suffit de remarquer que *le cube* I^n *est homéomorphe à la boule* $\mathbb{B}^n$. [On peut définir un homéomorphisme de $\mathbb{B}^n$ sur I^n par exemple à l'aide de la formule

$$\mathbf{x} \mapsto \lambda(\mathbf{x})\,\mathbf{x}, \quad \mathbf{x} \in \mathbb{B}^n,$$

où $\lambda(0) = 0$ et $\lambda(\mathbf{x})$ pour $\mathbf{x} \neq 0$ est la longueur du vecteur $\overrightarrow{OM}$ colinéaire au vecteur $\mathbf{x}$ et tel que le point M appartienne à la fron-



tière du cube I^n (il est aisé de voir que le vecteur $\overrightarrow{OM}$ est univoquement défini au signe près par cette condition).] □

Remarque 1. Le théorème 1 découle sans peine du théorème 2. En effet, s'il existait une rétraction $r: \mathbb{B}^n \to \mathbb{S}^{n-1}$, l'application composée

$$i \circ a \circ r: \mathbb{B}^n \to \mathbb{B}^n,$$

où $a: \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$ est *l'application antipodale* $\mathbf{x} \mapsto -\mathbf{x}$, et i l'injection $\mathbb{S}^{n-1} \to \mathbb{B}^n$, serait une application de $\mathbb{B}^n$ dans $\mathbb{B}^n$ sans point fixe.

* * *

Définition 2. Soient A et B des sous-ensembles disjoints d'un espace topologique $\mathscr{X}$. On dit qu'un ensemble fermé $C \subset \mathscr{X}$ *sépare* A de B (ou qu'il est une *cloison* entre A et B) si son complémentaire $\mathscr{X} \setminus C$ est la réunion d'ensembles U et V disjoints (et automatiquement ouverts) tels que $A \subset U$ et $B \subset V$.

La cloison C peut bien sûr être vide (ce sera le cas si l'espace $\mathscr{X}$ est la réunion d'ensembles U et V ouverts disjoints contenant respectivement les ensembles A et B).

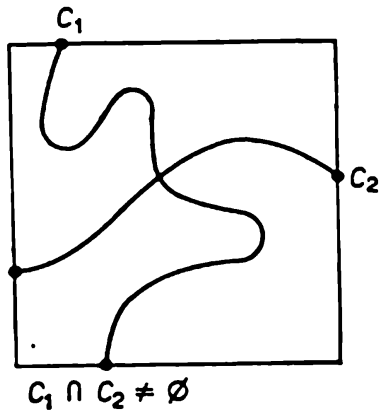
Pour tout $k = 1, \ldots, n$ et tout $\varepsilon = \pm 1$, désignons par $I_k^{n-1}(\varepsilon)$ la face du cube I^n composée des points $\mathbf{t} = (t_1, \ldots, t_n) \in I^n$ tels que $t_k = \varepsilon$.

Théorème 3 (des cloisons dans le cube). *L'intersection $C_1 \cap \ldots \cap C_n$ de toutes cloisons $C_1, \ldots, C_n$ entre les faces $I_1^{n-1}(-1), \ldots, I_n^{n-1}(-1)$ et les faces $I_1^{n-1}(+1), \ldots, I_n^{n-1}(+1)$ du cube I^n n'est pas vide:*

$$C_1 \cap \ldots \cap C_n \neq \varnothing.$$

D é m o n s t r a t i o n. Par hypothèse, pour tout $k = 1, \ldots, n$,

$$I^n \setminus C_k = U_k(-1) \cup U_k(+1),$$



où $U_k(-1)$ et $U_k(+1)$ sont des ensembles ouverts disjoints tels que $I_k^{n-1}(-1) \subset U_k(-1)$ et $I_k^{n-1}(+1) \subset U_k(+1)$. Pour tout point $\mathbf{t} \in I^n$ posons

$$\rho_k(\mathbf{t}) = \inf_{\mathbf{s} \in C_k} |\mathbf{t} - \mathbf{s}|, \quad k = 1, \ldots, n.$$

Remarquons qu'en raison de la compacité de l'ensemble C_k, l'égalité $\rho_k(\mathbf{t}) = 0$ équivaut à l'appartenance $\mathbf{t} \in C_k$.

Lemme 1. *Le nombre* $\rho_k(\mathbf{t})$ *est tel que* $|t_k - \varepsilon\rho_k(\mathbf{t})| \leqslant 1$, *pourvu que* $\mathbf{t} = (t_1, \ldots, t_n) \in U_k(\varepsilon)$.

De ce lemme il résulte que la formule

$$f(\mathbf{t}) = (t_1 - \tau_1, \ldots, t_n - \tau_n),$$

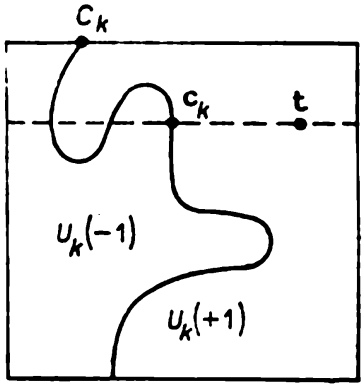
où

$$\tau_k = \begin{cases} \varepsilon\rho_k(\mathbf{t}) & \text{si } \mathbf{t} \in U_k(\varepsilon), \\ 0 & \text{si } \mathbf{t} \in C_k, \end{cases} \quad k = 1, \ldots, n,$$

définit de façon unique une application $f: I^n \to I^n$ (visiblement continue) telle que $f(\mathbf{t}) = \mathbf{t}$ si et seulement si $\rho_k(\mathbf{t}) = 0$ pour tout $k = 1, \ldots, n$, i.e. lorsque $\mathbf{t} \in C_1 \cap \ldots \cap C_n$. Puisque le corollaire 1 du théorème 2 nous dit qu'il existe dans le cube I^n un point $\mathbf{t}_0$ tel que $f(\mathbf{t}_0) = \mathbf{t}_0$, ceci prouve que $C_1 \cap \ldots \cap C_n \neq \varnothing$. □

Reste à prouver le lemme 1.

Démonstration du lemme 1. Supposons qu'un point $\mathbf{t} \in I^n$ appartient à un ensemble ouvert $U_k(\varepsilon)$. La cloison C_k



séparant $I_k^{n-1}(-1)$ de $I_k^{n-1}(+1)$, la perpendiculaire à ces faces passant par le point $\mathbf{t}$ coupe la cloison C_k en un point $\mathbf{c}_k$ compris entre le point $\mathbf{t}$ et sa projection sur la face $I_k^{n-1}(-\varepsilon)$. Si $\varepsilon = +1$, la coordonnée s_k du point $\mathbf{c}_k$ vérifie l'inégalité $s_k \leqslant t_k$, et si $\varepsilon = -1$, l'inégalité $s_k \geqslant t_k$. Ceci étant, $|s_k - t_k| = \sigma_k$, où σ_k est la distance

des points t et c_k. Donc,

$$s_k = t_k - \varepsilon\sigma_k.$$

Puisque par définition

$$\sigma_k \leqslant \rho_k(\mathrm{t}),$$

il s'ensuit que pour $\varepsilon = +1$

$$t_k - \varepsilon\rho_k(\mathrm{t}) = t_k - \rho_k(\mathrm{t}) \geqslant t_k - \sigma_k = t_k - \varepsilon\sigma_k = s_k \geqslant -1,$$

et pour $\varepsilon = -1$

$$t_k - \varepsilon\rho_k(\mathrm{t}) = t_k + \rho_k(\mathrm{t}) \leqslant t_k + \sigma_k = t_k - \varepsilon\sigma_k = s_k \leqslant 1.$$

D'autre part, pour $\varepsilon = +1$

$$t_k - \varepsilon\rho_k(\mathrm{t}) = t_k - \rho_k(\mathrm{t}) \leqslant t_k \leqslant 1,$$

et pour $\varepsilon = -1$

$$t_k - \varepsilon\rho_k(\mathrm{t}) = t_k + \rho_k(\mathrm{t}) \geqslant t_k \geqslant -1.$$

Donc, dans tous les cas

$$|\, t_k - \varepsilon\rho_k(\mathrm{t}) \,| \leqslant 1. \quad \square$$

Remarque 2. Le théorème 1 résulte aussi aisément du théorème 3. En effet, soit $\dot{I}^n$ la *frontière* du cube I^n (qui est composée des points $\mathrm{t} \in I^n$ tels que $t_k = \pm 1$ au moins pour un $k = 1, \ldots, n$) et soit S_k, $k = 1, \ldots, n$, le sous-ensemble de $\dot{I}^n$ composé des points t tels que $t_k = 0$. Il est clair que S_k est une cloison dans $\dot{I}^n$ entre les faces $I_k^{n-1}(-1)$ et $I_k^{n-1}(+1)$. Donc, pour toute rétraction $r : I^n \to \dot{I}^n$, l'image réciproque $C_k = r^{-1}S_k$ de l'ensemble S_k sera une cloison entre les images réciproques $r^{-1}I_k^{n-1}(-1)$ et $r^{-1}I_k^{n-1}(+1)$ des faces $I_k^{n-1}(-1)$ et $I_k^{n-1}(+1)$, donc — puisque $I_k^{n-1}(\varepsilon) \subset r^{-1}I_k^{n-1}(\varepsilon)$ — entre les faces $I_k^{n-1}(-1)$ et $I_k^{n-1}(+1)$. D'autre part, puisque de toute évidence $S_1 \cap \ldots \cap S_n = \varnothing$, il s'ensuit que $C_1 \cap \ldots \cap C_n = \varnothing$, ce qui est impossible d'après le théorème 3. Donc, il ne peut exister de rétraction r pour le couple $(I^n, \dot{I}^n)$ et, par conséquent, pour le couple homéomorphe $(\mathbb{B}^n, \mathbb{S}^{n-1})$. $\square$

* * *

Les cloisons n'existent pas dans tous les espaces, ce qui nous contraint à introduire la définition suivante.

Définition 3. On dit qu'un espace $\mathscr{X}$ est *normal* s'il existe une cloison pour deux quelconques de ses sous-ensembles A et B fermés disjoints, et *complètement normal* s'il existe une cloison pour tous

sous-ensembles A et B tels que l'adhérence de l'un d'eux ne rencontre pas l'autre (de tels ensembles A et B sont dits *séparés*).

Remarquons que la séparabilité des ensembles est une condition nécessaire d'existence d'une cloison entre A et B.

Proposition 1. *Un espace normal $\mathscr{X}$ est complètement normal si et seulement si chacun de ses sous-espaces est normal.*

D é m o n s t r a t i o n. Il est clair que chaque sous-espace d'un espace complètement normal est complètement normal (car des ensembles séparés dans un sous-espace le sont dans l'espace tout entier), donc est normal. Réciproquement, soient A et B des sous-ensembles séparés de $\mathscr{X}$ et supposons que

$$O = \mathscr{X} \setminus (\bar{A} \cap \bar{B}) = (\mathscr{X} \setminus \bar{A}) \cup (\mathscr{X} \setminus \bar{B}).$$

Les ensembles

$$\bar{A} \cap O = \bar{A} \setminus (\bar{A} \cap \bar{B}) \quad \text{et} \quad \bar{B} \cap O = \bar{B} \setminus (\bar{A} \cap \bar{B})$$

sont fermés dans O, donc, si O est normal, ils sont séparés par une cloison C. Par définition, $O \setminus C = U \cup V$, où U et V sont des ensembles ouverts dans O disjoints tels que $\bar{A} \cap O \subset U$ et $\bar{B} \cap O \subset V$. Mais puisque O est visiblement ouvert dans $\mathscr{X}$, il en est de même de U et de V, et comme $A \cap \bar{B} = \varnothing$ et $B \cap \bar{A} = \varnothing$ (i.e. A et B sont séparés), il vient que $A \subset \bar{A} \cap O$ et $B \subset \bar{B} \cap O$. Donc, $A \subset U$ et $B \subset V$. Puisque

$$O \setminus C = \mathscr{X} \setminus (C \cup (\bar{A} \cap \bar{B})),$$

ceci prouve que l'ensemble $C \cup (\bar{A} \cap \bar{B})$ est une cloison dans $\mathscr{X}$ entre A et B. Donc, un espace normal, dont tout sous-espace est aussi normal, est complètement normal. □

Aux termes de la proposition 1, les espaces complètement normaux s'appellent aussi *héréditairement normaux*.

Remarque 3. Dans la définition des espaces normaux (et complètement normaux), à la condition d'existence des cloisons on ajoute souvent la condition de séparabilité de l'espace. Signalons que pour notre part nous ne le ferons pas.

Proposition 2. *Tout espace métrique est normal.*

D é m o n s t r a t i o n. Soient A et B des sous-espaces fermés disjoints d'un espace métrique $\mathscr{X}$. Pour tout point $p \in A$ le nombre

$$\rho_B(p) = \inf_{q \in B} \rho(p, q)$$

(*distance de p à B*) est de toute évidence strictement positif et donc, pour ce point, est défini le $\rho_B(p)/2$-voisinage sphérique

$$U(p) = \left\{a \in \mathscr{X},\ \rho(p, a) < \frac{\rho_B(p)}{2}\right\}.$$

De façon analogue, pour tout point $q \in B$, est défini le $\rho_A(q)/2$-voisinage sphérique

$$V(q) = \left\{a \in \mathscr{X},\ \rho(a,\ q) < \frac{\rho_A(q)}{2}\right\},$$

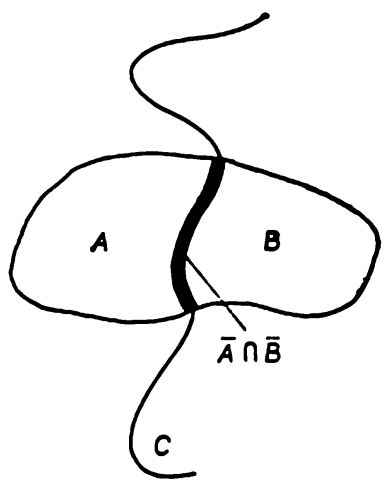
où

$$\rho_A(q) = \inf_{p \in A} \rho(p,\ q)$$

est la distance de q à A. Soient

$$U = \bigcup_{p \in A} U(p), \quad V = \bigcup_{q \in B} V(q).$$

Les ensembles U et V sont ouverts, contiennent respectivement les ensembles A et B et sont disjoints (si $a \in U \cap V$, il existe alors



des points $p \in A$, $q \in B$ tels que $a \in U(p)$, $a \in V(q)$, et donc

$$\rho(p,\ a) < \frac{\rho_B(p)}{2} \leqslant \frac{\rho(p,\ q)}{2}$$

et

$$\rho(a,\ q) < \frac{\rho_A(q)}{2} \leqslant \frac{\rho(p,\ q)}{2};$$

mais alors

$$\rho(p,\ q) \leqslant \rho(p,\ a) + \rho(a,\ q) < \rho(p,\ q),$$

ce qui est impossible). Donc, l'ensemble $C = \mathscr{X} \setminus (U \cup V)$ est une cloison entre A et B. □

Corollaire. *Tout espace métrique est complètement normal.*

Démonstration. Il suffit de remarquer que tout sous-espace d'un espace métrique est un espace métrique (donc, est normal). □

En particulier, *le cube I^n est un espace complètement normal.*

Proposition 3. *Tout espace séparé compact $\mathscr{X}$ est normal.*

Démonstration. Soient A et B des sous-ensembles fermés disjoints de l'espace $\mathscr{X}$. L'espace $\mathscr{X}$ étant séparé, pour tout point $p \in A$ et tout point $q \in B$ il existe un voisinage $U_q(p)$ de p et un voisinage $V_p(q)$ de q tels que

$$U_q(p) \cap V_p(q) = \varnothing.$$

Pour tout point $q \in B$, les voisinages $U_q(p)$, $p \in A$, forment un recouvrement ouvert du sous-espace A et, par suite — puisque ce sous-espace est compact en tant qu'ensemble fermé d'un espace compact — il existe une famille finie $\{p_1, \ldots, p_n\}$ de points de A, telle que l'ensemble ouvert

$$U_q = U_q(p_1) \cup \ldots \cup U_q(p_n)$$

contient A. Ceci étant, l'ensemble U_q ne rencontre pas le voisinage

$$V(q) = V_{p_1}(q) \cap \ldots \cap V_{p_n}(q)$$

du point q.

Les voisinages $V(q)$, $q \in B$, forment un recouvrement ouvert du sous-espace B et donc, pour les mêmes raisons, il existe une famille de points $q_1, \ldots, q_m$ de l'ensemble B, telle que l'ensemble ouvert

$$V = V(q_1) \cup \ldots \cup V(q_m)$$

contient V. Ceci étant, l'ensemble V ne rencontre pas l'ensemble ouvert

$$U = U_{q_1} \cap \ldots \cap U_{q_m},$$

qui contient A. Donc, l'ensemble $C = \mathscr{X} \setminus (U \cup V)$ est une cloison entre A et B. □

Nous citerons à la prochaine leçon un exemple d'espace compact séparé qui n'est pas complètement normal.

* * *

Nous aurons besoin aussi de la proposition simple suivante.

Proposition 4 (de prolongement des cloisons). *Soient $\mathscr{X}$ un espace complètement normal, A et B des sous-ensembles fermés disjoints de $\mathscr{X}$ et $\mathscr{Y}$ un sous-espace fermé de $\mathscr{X}$. Soit par ailleurs C' une cloison dans $\mathscr{Y}$ séparant les ensembles $A \cap \mathscr{Y}$ et $B \cap \mathscr{Y}$ fermés (dans $\mathscr{Y}$). Il existe alors dans $\mathscr{X}$ une cloison C entre A et B telle que $C \cap \mathscr{Y} \subset C'$.*

Démonstration. Par hypothèse, $\mathscr{Y} \setminus C' = U' \cup V'$, où U' et V' sont des ensembles disjoints ouverts (dans $\mathscr{Y}$) contenant respectivement les ensembles $A \cap \mathscr{Y}$ et $B \cap \mathscr{Y}$. Soient

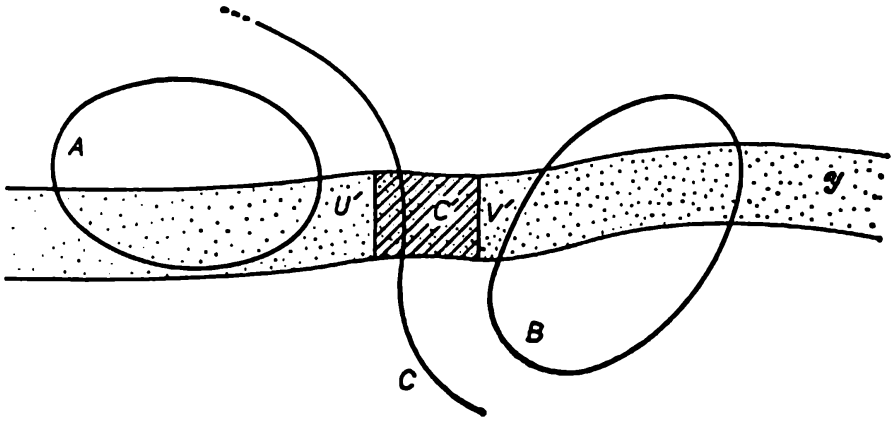
$$A' = A \cup U' \quad \text{et} \quad B' = B \cup V'.$$

Les ensembles U' et V' étant non seulement ouverts mais aussi fermés (dans $\mathcal{Y} \setminus C'$), ils sont séparés dans $\mathcal{X}$. Donc, il en est de même des ensembles A' et B', et il existe par conséquent une cloison C entre A et B. Par ailleurs, puisque

$$\mathcal{X} \setminus C \supset A' \cup B' \supset U' \cup V' = \mathcal{Y} \setminus C',$$

alors $\mathcal{Y} \cap C \subset C'$. □

Nous pouvons maintenant reprendre l'étude des cloisons dans le cube I^n.



Proposition 5. *Soit $Q_1 \supset Q_2 \supset \ldots \supset Q_n$ une suite strictement décroissante de sous-ensembles fermés du cube I^n telle que*

Q_1 *sépare* $I_1^{n-1}(-1)$ *de* $I_1^{n-1}(+1)$ *dans* I^n;

Q_2 *sépare* $I_2^{n-1}(-1) \cap Q_1$ *de* $I_2^{n-1}(+1) \cap Q_1$ *dans* Q_1;

. . . . . .

Q_n *sépare* $I_n^{n-1}(-1) \cap Q_{n-1}$ *de* $I_n^{n-1}(+1) \cap Q_{n-1}$ *dans* Q_{n-1}.

Alors Q_n n'est pas vide.

Démonstration. En appliquant la proposition 4 à $\mathcal{X} = I^n$, $A = I_k^{n-1}(-1)$, $B = I_k^{n-1}(+1)$ et $\mathcal{Y} = Q_{k-1}$, on obtient pour tout $k = 2, \ldots, n$ dans le cube I^n une cloison C_k séparant $I_k^{n-1}(-1)$ et $I_k^{n-1}(+1)$, telle que $C_k \cap Q_{k-1} \subset Q_k$. Pour $k = 1$, on posera $C_1 = Q_1$.

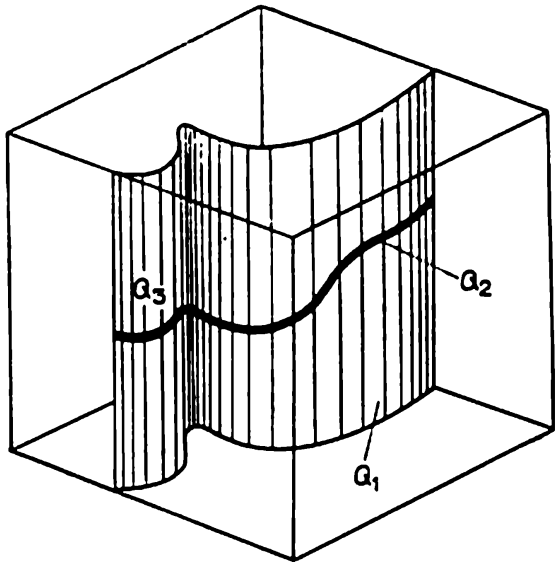
D'après le théorème 3, les cloisons $C_1, C_2, \ldots, C_n$ sont telles que

$$C_1 \cap C_2 \cap \ldots \cap C_n \neq \varnothing.$$

D'autre part, puisque

$$C_1 = Q_1, \quad C_2 \cap Q_1 \subset Q_2, \ \ldots, \ C_n \cap Q_{n-1} \subset Q_n,$$

on a $C_1 \cap C_2 \cap \ldots \cap C_n \subset Q$. Donc, $Q_n \neq \varnothing$. □



* * *

Nous pouvons maintenant prouver le théorème fondamental de cette leçon.

Théorème 4 (de Lebesgue de recouvrement du cube). *Tout recouvrement fermé fini $\{C_\alpha\}$ du cube I^n, composé d'ensembles dont aucun ne rencontre deux faces quelconques opposées, est de multiplicité $\geqslant n+1$.*

Démonstration. Soient

A_1 la réunion de tous les éléments du recouvrement $\{C_\alpha\}$ rencontrant la face $I_1^{n-1}(-1)$ (donc, pas la face $I_1^{n-1}(+1)$);

A_2 la réunion de tous les éléments du recouvrement $\{C_\alpha\}$ rencontrant la face $I_2^{n-1}(-1)$ et n'intervenant pas dans A_1;

. . . . . .

A_n la réunion de tous les éléments du recouvrement $\{C_\alpha\}$ rencontrant la face $I_n^{n-1}(-1)$ et n'intervenant ni dans A_1, ni dans $A_2, \ldots$, ni dans A_{n-1};

A_{n+1} la réunion de tous les éléments du recouvrement $\{C_\alpha\}$ ne rencontrant aucune face du cube I^n.

Supposons par ailleurs que

$$Q_1 = A_1 \cap B_1, \text{ où } B_1 = A_2 \cup \ldots \cup A_{n+1},$$
$$Q_2 = A_1 \cap A_2 \cap B_2, \text{ où } B_2 = A_3 \cup \ldots \cup A_{n+1},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$Q_{n-1} = A_1 \cap A_2 \cap \ldots \cap A_{n-1} \cap B_{n-1}, \text{ où } B_{n-1} = A_n \cup A_{n+1},$$
$$Q_n = A_1 \cap A_2 \cap \ldots \cap A_n \cap B_n, \text{ où } B_n = A_{n+1}.$$

Il est clair que les ensembles Q_k sont fermés et forment une suite strictement décroissante

$$Q_1 \supset Q_2 \supset \ldots \supset Q_n.$$

De plus, puisque $A_1 \cup B_1 = I^n$, on a $(I^n \setminus A_1) \cap (I^n \setminus B_1) = \varnothing$, et comme l'ensemble A_1 ne rencontre pas la face $I_1^{n-1}(+1)$, et les ensembles $A_2, \ldots, A_{n+1}$, la face $I_1^{n-1}(-1)$, il s'ensuit que

$$I_1^{n-1}(+1) \subset I^n \setminus A_1 \quad \text{et} \quad I_1^{n-1}(-1) \subset I^n \setminus B_1.$$

Vu que $I^n \setminus Q_1 = (I^n \setminus A_1) \cup (I^n \setminus B_1)$, ceci prouve que Q_1 est une cloison entre $I_1^{n-1}(-1)$ et $I_1^{n-1}(+1)$ dans I^n.

De façon analogue, puisque $(A_1 \cap A_2) \cup (A_1 \cap B_2) = A_1 \cap \cap (A_2 \cup B_2) = A_1 \cap B_1 = Q_1$, alors $(Q \setminus (A_1 \cap A_2)) \cap (Q_1 \setminus (A_1 \cap \cap B_2)) = \varnothing$, et comme l'ensemble A_2 ne rencontre pas la face $I_2^{n-1}(+1)$, et les ensembles $A_3, \ldots, A_{n+1}$, la face $I_2^{n-1}(-1)$, on déduit que

$$I_2^{n-1}(+1) \cap Q_1 \subset Q_1 \setminus (A_1 \cap A_2)$$

et

$$I_2^{n-1}(-1) \cap Q_1 \subset Q_1 \setminus (A_1 \cap B_2).$$

Comme $Q_2 = (A_1 \cap A_2) \cap (A_1 \cap B_2)$ et donc

$$Q_1 \setminus Q_2 = (Q_1 \setminus (A_1 \cap A_2)) \cup (Q_1 \setminus (A_1 \cap B_2)),$$

ceci prouve que Q_2 est une cloison entre $I_2^{n-1}(-1) \cap Q_1$ et $I_2^{n-1}(+1) \cap \cap Q_1$ dans Q_1.

D'une façon générale, si $A_k' = A_1 \cap A_2 \cap \ldots \cap A_{k+1}$ et $B_k' = = A_1 \cap A_2 \cap \ldots \cap A_k \cap B_{k+1}$, où $k = 1, \ldots, n-1$, alors

$$A_k' \cup B_k' = A_1 \cap \ldots \cap A_k \cap (A_{k+1} \cup B_{k+1}) = \\ = A_1 \cap \ldots \cap A_k \cap B_k = Q_k,$$

donc $(Q_k \setminus A_k') \cap (Q_k \setminus B_k') = \varnothing$. D'autre part, puisque l'ensemble A_{k+1}, donc l'ensemble A_k' ne rencontrent pas la face $I_{k+1}^{n-1}(+1)$, et les ensembles $A_{k+2}, \ldots, A_{n+1}$, donc les ensembles B_k et B_k', ne rencontrent pas la face $I_{k+1}^{n-1}(-1)$, il vient

$$I_{k+1}^{n-1}(+1) \cap Q_k \subset Q_k \setminus A_k' \quad \text{et} \quad I_{k+1}^{n-1}(-1) \cap Q_k \subset Q_k \setminus B_k'.$$

Comme $Q_{k+1} = A_k' \cap B_k'$, donc $Q_k \setminus Q_{k+1} = (Q_k \setminus A_k') \cup (Q_k \setminus B_k')$, ceci prouve que Q_{k+1} est une cloison entre $I_{k+1}^{n-1}(-1) \cap Q_k$ et $I_{k+1}^{n-1}(+1) \cap Q_k$ dans Q_k.

Donc, les ensembles $Q_1, \ldots, Q_n$ satisfont à toutes les conditions de la proposition 5 et, par conséquent, en vertu de cette même proposition, l'ensemble Q_n n'est pas vide.

Soit p_0 un point de Q_n. Puisque

$$Q_n = A_1 \cap \ldots \cap A_{n+1}$$

et que chaque ensemble A_i est la réunion d'éléments du recouvrement $\{C_\alpha\}$ (les éléments intervenant dans un A_k n'interviennent dans aucun autre), le point p_0 appartient au moins à $n+1$ de ces éléments. Donc, le recouvrement $\{C_\alpha\}$ est de multiplicité $\geqslant n+1$. $\square$

* * *

Nous sommes tout près de déduire l'inégalité dim $I^n \geqslant n$ du théorème de Lebesgue.

Définition 4. Un recouvrement ouvert $\{U_\alpha\}$ d'un espace topologique $\mathscr{X}$ est dit *coercible* s'il existe un recouvrement ouvert $\{V_\alpha\}$ de l'espace $\mathscr{X}$ (appelé *coercition* de $\{U_\alpha\}$) tel que

$$\bar{V}_\alpha \subset U_\alpha \text{ pour tout } \alpha.$$

Proposition 6. *Tout recouvrement ouvert d'un espace normal $\mathscr{X}$ est coercible.*

Démonstration. Nous allons démontrer cette proposition sous l'hypothèse que *le recouvrement est fini.* Le cas général implique une récurrence transfinie et ne nous sert à rien ici.

Soit $\{U_1, \ldots, U_n\}$ un recouvrement ouvert fini d'un espace normal $\mathscr{X}$. Il suffit de toute évidence de prouver qu'il existe un ensemble ouvert V_1 tel que $\bar{V}_1 \subset U_1$ et que la famille $\{V_1, U_2, \ldots, U_n\}$ soit encore un recouvrement.

A cet effet considérons les ensembles fermés

$$\mathscr{X} \setminus U_1 \quad \text{et} \quad \mathscr{X} \setminus (U_2 \cup \ldots \cup U_n).$$

Ces ensembles sont disjoints, puisque $\{U_1, U_2, \ldots, U_n\}$ est un recouvrement. Il existe donc une cloison C qui les sépare. Soit $\mathscr{X} \setminus C = U \cup V$, où U et V sont des ensembles ouverts disjoints tels que

$$\mathscr{X} \setminus U_1 \subset U \quad \text{et} \quad \mathscr{X} \setminus (U_2 \cup \ldots \cup U_n) \subset V.$$

La deuxième inclusion exprime que la famille $\{V, U_2, \ldots, U_n\}$ est un recouvrement de $\mathscr{X}$, et la première entraîne $\bar{V} \subset \mathscr{X} \setminus U \subset U_1$. Reste à poser $V_1 = V$. □

Nous pouvons prouver maintenant l'inégalité dim $I^n \geqslant n$. Le cube I^n étant un espace métrique compact, cette inégalité exprime, en vertu du corollaire 2 du lemme 1 de la leçon précédente, que pour un $\varepsilon > 0$ (on montrera que tout $\varepsilon < 2$ convient) tout ε-recouvrement ouvert fini $\{U_\alpha\}$ du cube I^n est de multiplicité $\geqslant n + 1$.

Soit $\{V_\alpha\}$ une coercition quelconque du recouvrement $\{U_\alpha\}$ (cette coercition existe pour les recouvrements finis en vertu de la proposition 3 prouvée ci-dessus). La famille $\{\bar{V}_\alpha\}$ sera alors un ε-recouvrement fermé du cube I^n. Puisque pour $\varepsilon < 2$ ce recouvrement satisfait visiblement à l'hypothèse du théorème de Lebesgue, la multiplicité du recouvrement $\{\bar{V}_\alpha\}$ sera $\geqslant n + 1$ d'après ce théorème. Mais alors le recouvrement $\{U_\alpha\}$ sera aussi visiblement de multiplicité $\geqslant n + 1$. □

LEÇON 10

Ordinaux. — Topologie de l'ordre dans les ensembles d'ordinaux. — Espaces de dimension 0. — Exemple de Tykhonov. — Produit tykhonovien d'espaces topologiques. — Filtres. — Ensembles centrés d'ensembles. — Ultrafiltres. — Critère de compacité. — Théorème de Tykhonov.

Cette leçon, qui est entièrement consacrée à la topologie générale et qui n'est pas rattachée à la théorie des variétés différentiables, est composée de deux parties liées au seul nom de Tykhonov. Dans la première, on construit un exemple de Tykhonov d'espace séparé compact de dimension 0, normal mais non complètement normal, contenant un sous-espace de dimension un, et dans la deuxième, on prouve le théorème de Tykhonov du produit d'espaces compacts.

* * *

L'exemple de Tykhonov se construit à l'aide des ordinaux transfinis. Rappelons brièvement leur définition et leurs propriétés fondamentales.

On dit qu'un ensemble A est *ordonné* s'il est muni d'une relation $\leqslant$ *réflexive* ($a \leqslant a$ pour tout $a \in A$), *antisymétrique* ($a \leqslant b$ et $b \leqslant a$ impliquent $a = b$, et *transitive* ($a \leqslant b$ et $b \leqslant c$ entraînent $a \leqslant c$). Si $a \leqslant b$, mais $a \neq b$, on écrit $a < b$. Une application $\varphi: A \to B$ d'ensembles ordonnés est dite *monotone* si $\varphi a \leqslant \varphi b$ pour $a \leqslant b$. Une application monotone bijective dont la réciproque est aussi monotone s'appelle *isomorphisme*. Des ensembles ordonnés A et B sont dits *isomorphes* s'il existe au moins un isomorphisme de A sur B. La relation d'isomorphie d'ensembles ordonnés est une relation d'équivalence et donc tous les ensembles ordonnés se répartissent en classes d'ensembles isomorphes.

Un ensemble ordonné A est dit *totalement ordonné* si deux quelconques de ses éléments a et b sont comparables (i.e. soit $a \leqslant b$, soit $b \leqslant a$). Dans le cas contraire, il est dit *partiellement ordonné*. Les classes d'ensembles totalement ordonnés et isomorphes sont dites *de même type d'ordre*. Une application d'ensembles totalement ordonnés est un isomorphisme si et seulement si elle est bijective et strictement monotone.

Un ensemble A totalement ordonné est *bien ordonné* si toute partie non vide de A (y compris l'ensemble A) admet un premier élément. Les types d'ordre des ensembles bien ordonnés s'appellent *ordinaux*. Si un ensemble bien ordonné A est associé à un ordinal α, on dit qu'il est *de type* α. On dit aussi que α *est l'ordinal de l'ensemble* A.

Si des ensembles totalement ordonnés A et B sont de même type α, ils sont de toute évidence équipotents. Leur cardinal s'appelle *cardinal du nombre* α et se désigne par $|\alpha|$.

Si un ensemble A est fini, son ordinal est dit *fini* (ou *de première classe*). Puisque pour tout entier naturel n deux ensembles totalement ordonnés quel-

conques de cardinal n sont visiblement isomorphes, l'égalité $|\alpha| = |\beta|$, où α et β sont des ordinaux finis, entraîne $\alpha = \beta$. Donc, ces nombres peuvent être identifiés à leurs cardinaux, i.e. avec les entiers naturels

$$0, 1, 2, \ldots, n, \ldots \tag{1}$$

(Les entiers naturels conviennent aussi bien pour le calcul que pour l'énumération.)

Si un ensemble A est dénombrable, son ordinal α est dit *dénombrable* (ou *de deuxième classe*). Un exemple d'ordinal dénombrable est le type d'ordre de l'ensemble (1) des entiers naturels (qui est visiblement un ensemble bien ordonné). Ce nombre est désigné par ω.

Pout tout élément a_0 d'un ensemble bien ordonné A, l'ensemble de tous les éléments $a \in A$ tels que $a < a_0$ s'appelle *segment* (ou *intervalle initial*) de l'ensemble A défini par a_0 et se note par le symbole $A(a_0)$. L'ensemble $A(a_0)$ est lui aussi bien ordonné.

Si tout ensemble de type β est isomorphe à un segment d'un ensemble de type α, on dit que β est *strictement inférieur* à α et on écrit $\beta < \alpha$. La formule $\beta \leqslant \alpha$ signifie par définition que soit $\beta < \alpha$, soit $\beta = \alpha$.

Il est clair que *la relation* $<$ *est transitive* ($\alpha < \beta$ et $\beta < \gamma$ impliquent $\alpha < \gamma$).

L'ensemble de tous les ordinaux β tels que $\beta < \alpha$ se note $[0, \alpha[$ et s'appelle *intervalle d'extrémité* α. De façon analogue, l'ensemble de tous les ordinaux β tels que $\beta \leqslant \alpha$ se note $[0, \alpha]$ et s'appelle *segment d'extrémité* α. Il est clair que $[0, \alpha] = [0, \alpha[\cup \{\alpha\}$.

Par exemple, l'intervalle $[0, \omega[$ est l'ensemble (1) de tous les entiers naturels, et le segment $[0, \omega]$, l'ensemble (1) auquel on a ajouté l'élément ω à droite.

Puisque tout segment de l'ensemble (1) est de la forme $[0, n]$, on voit en particulier que ω *est le plus petit ordinal de deuxième classe.*

Il est aisé de voir qu'*aucun ensemble bien ordonné A ne peut être isomorphe à un segment $B(a_0)$ de l'un de ses sous-ensembles B (le cas $B = A$ n'est pas exclu)*. En effet, supposons par absurde qu'il existe un isomorphisme $\varphi: A \to B(a_0)$ et considérons l'ensemble de tous les éléments $a \in A$ tels que $\varphi a < a$. Cet ensemble contient a_0 et donc n'est pas vide. Soit a_1 son premier élément (cet élément existe, car l'ensemble A est bien ordonné) et soit $a_2 = \varphi a_1$. Alors $a_2 < a_1$ et en même temps $\varphi a_2 < \varphi a_1 = a_2$, ce qui contredit le choix de a_1. Cette contradiction prouve que l'isomorphisme φ ne peut pas exister. □

Pour $B = A$, on constate en particulier que *si* $\beta < \alpha$, *on a* $\beta \neq \alpha$.

Il est aisé de voir aussi que *tout intervalle* $[0, \alpha[$ *est un ensemble bien ordonné de type d'ordre* α. En effet, en associant à tout élément a d'un ensemble bien ordonné A de type d'ordre α le type d'ordre du segment $A(a)$, on obtient de toute évidence une application isomorphe $A \to [0, \alpha[$. □

Il s'ensuit que *chaque ensemble d'ordinaux A possède un premier élément.* En effet, soit α_0 un élément quelconque de A. Si α_0 est le premier élément, il n'y a rien à prouver; si α_0 ne l'est pas, l'intersection $[0, \alpha_0[\cap A$ n'est pas vide et le premier élément de cette intersection (lequel existe puisque l'intervalle $[0, \alpha_0[$ est bien ordonné) sera le premier élément de A. □

L'intervalle $[0, \alpha[$ étant bien ordonné, il en sera de même du segment $[0, \alpha] = [0, \alpha[\cup \{\alpha\}$. Désignons le type d'ordre de ce segment par $\alpha + 1$. L'intervalle $[0, \alpha[$ étant l'intervalle initial $[0, \alpha]$ (α) du segment $[0, \alpha]$, on a $\alpha < \alpha + 1$. Par ailleurs, *si* $\beta \leqslant \alpha + 1$, *alors* $\beta \leqslant \alpha$ (car tout intervalle initial du segment $[0, \alpha]$ est contenu dans l'intervalle $[0, \alpha[$). Pour cette raison on dit que le nombre $\alpha + 1$ *suit immédiatement* le nombre α.

Un ordinal β s'appelle *limite* s'il n'existe pas d'ordinal α tel que $\beta = \alpha + 1$. Le nombre ω est un exemple d'ordinal limite.

Soient α et β des ordinaux et soit $D = [0, \alpha[\cap [0, \beta[$ l'intersection des intervalles $[0, \alpha[$ et $[0, \beta[$. *Il existe un ordinal* $\delta \leqslant \alpha$ *tel que* $D = [0, \delta[$. En

effet, si $D = [0, \alpha[$, alors $\delta = \alpha$. Supposons que $D \neq [0, \alpha[$, i.e. que le complémentaire $C = [0, \alpha[\setminus D$ de l'ensemble D dans $[0, \alpha[$ n'est pas vide, et soit δ le premier élément de C. Si $\xi \in [0, \delta[$, alors $\xi < \alpha$ (car $\xi < \delta$ et $\delta < \alpha$) et $\xi \notin C$ (puisque $\xi < \delta$ et δ est le premier élément de C), i.e. $\xi \in D$. Réciproquement, si $\xi \in D$ et $\delta \leqslant \xi$, alors $\delta < \beta$ (car $\xi < \beta$) et, par suite, $\delta \in D$, ce qui est impossible. Donc, $\xi < \delta$, i.e. $\xi \in [0, \delta[$. □

Par raison de symétrie, on a simultanément $\delta \leqslant \alpha$ et $\delta \leqslant \beta$. Ceci étant, si $\delta < \alpha$ et $\delta < \beta$, alors $\delta \in D = [0, \delta[$, ce qui est impossible. Donc, soit $\delta = \alpha$ (et alors $\alpha \leqslant \beta$), soit $\delta = \beta$ (et alors $\beta \leqslant \alpha$). Ceci prouve que pour deux ordinaux quelconques α et β, soit $\alpha = \beta$, soit $\alpha < \beta$, soit $\beta < \alpha$, i.e. *chaque ensemble d'ordinaux est totalement ordonné, donc* (puisque chacun de ses sous-ensembles admet un premier élément), *bien ordonné.*

Il est aisé de voir maintenant que *le type β d'un sous-ensemble B d'un ensemble totalement ordonné A est $\leqslant$ au type α de A* :

$$\beta \leqslant \alpha.$$

En effet, dans le cas contraire on aurait $\alpha < \beta$ et l'ensemble A serait isomorphe à un segment du sous-ensemble B, ce qui est impossible. □

Soit donnée une famille $\{\alpha_\xi\}$ d'ordinaux indexés par les nombres ξ d'un intervalle $[0, \tau[$, et soit A_ξ un ensemble de type α_ξ. Ordonnons la réunion disjointe

$$A = \bigsqcup_\xi A_\xi$$

de tous les ensembles A_ξ en convenant que pour des éléments $a \in A_\xi$ et $b \in A_\eta$ l'inégalité $a < b$ a lieu si et seulement si soit $\xi < \eta$, soit $\xi = \eta$ et $a < b$ dans l'ensemble A_ξ. Pour toute partie non vide $C \subset A$ l'ensemble de tous les nombres $\xi < \tau$ tels que $C \cap A_\xi \neq \varnothing$ est une partie non vide de l'intervalle $[0, \tau[$, donc, admet un premier élément ξ_0. Il est clair que le premier élément de l'ensemble $C \cap A_{\xi_0}$ sera premier élément de l'ensemble C. Ceci prouve que *A est bien ordonné* et donc que son type d'ordre α est un ordinal. Il est évident que le nombre α ne dépend que des nombres α_ξ (et pas du choix des ensembles A_ξ). On l'appelle *somme* des nombres α_ξ et on le désigne par le symbole

$$\sum_{\xi<\tau} \alpha_\xi.$$

Signalons que α *dépend* de l'ordre de numérotation des nombres α_ξ. (Par exemple, étant donné qu'un ensemble de type $\omega + n$ admet un dernier élément et qu'un ensemble de type $n + \omega$ n'en admet pas, on a $n + \omega \neq \omega + n$.)

Puisque $A_\xi \subset A$ pour tout ξ, il vient $\alpha_\xi \leqslant \alpha$. (L'égalité est possible ici ; il est clair par exemple que $n + \omega = \omega$.)

Mais $\alpha_\xi < \alpha + 1$. Puisque tout ensemble d'ordinaux est d'un certain type τ en tant qu'ensemble bien ordonné et donc que ses éléments peuvent être numérotés avec les nombres de l'intervalle $[0, \tau[$, on a prouvé que *pour tout ensemble d'ordinaux il existe un nombre qui est strictement supérieur à tous ces ordinaux.*

[Il s'ensuit que *la notion d'ensemble de tous les ordinaux n'a pas de sens.* Le lecteur attentif aura certainement remarqué qu'au prix d'énoncés boiteux nous avons soigneusement évité de mentionner cet « ensemble ». Le lecteur a donc pris acte.]

Il existe en particulier des ordinaux strictement supérieurs à tous les nombres de deuxième classe (les nombres *de troisième classe*). Le plus petit de ces nombres, qui est le *premier nombre de troisième classe*, est désigné par Ω.

La somme $\sum_{n<\omega} \alpha_n$ de toute suite dénombrable (ou finie) $\{\alpha_n\}$ de nombres de deuxième classe est de toute évidence un nombre de deuxième classe (car la

réunion disjointe d'ensembles dénombrables est dénombrable). Donc, le plus petit des nombres strictement supérieurs à α (ce nombre est désigné par $\sup \alpha_n$) est aussi un nombre de deuxième classe (en bref, *si* $\alpha_n < \Omega$, *alors* $\sup \alpha_n < \Omega$).

L'existence du nombre Ω prouve qu'il existe des ordinaux non dénombrables. En fait *il existe des ordinaux de cardinal quelconque*, i.e., ce qui est visiblement équivalent, *tout ensemble A peut être numéroté par des ordinaux d'un intervalle* $[0, \alpha[$. Pour prouver cette proposition (connue sous le nom de théorème de Zermelo de bonne ordonnance), on dira qu'un ordinal β est *distingué* s'il existe un sous-ensemble de l'ensemble A susceptible d'être numéroté par les nombres de l'intervalle $[0, \beta[$. L'ensemble B de tous les nombres distingués n'est pas vide (si par exemple l'ensemble A est infini, tous les nombres de première et de deuxième classe sont distingués) et possède la propriété suivante: si $\beta \in B$ et $\gamma < \beta$, alors $\gamma \in B$, i.e. $[0, \beta[\subset B$.

Soit $*$ un élément n'appartenant pas à A et soit $A^* = A \sqcup \{*\}$. Construisons une application $\varphi: B \to A^*$ de la manière suivante:

a) Choisissons un élément arbitraire $a_0 \in A$ et posons $\varphi(0) = a_0$.

b) Supposons que l'application φ a déjà été construite pour un nombre $\beta \in B$ sur l'intervalle $[0, \beta[$ et soit $A' = \varphi[0, \beta[$. Si $A' \subset A$ et $A' \neq A$, choisissons arbitrairement un élément $a_\beta \in A \setminus A'$ et posons

$$\varphi(\beta) = a_\beta.$$

Dans le cas contraire (i.e. si ou bien $A' = A$, ou bien $* \in A'$) nous poserons $\varphi(\beta) = *$.

Il est évident que cette construction définit φ sur B tout entier. [L'ensemble B' de tous les $\beta \in B$ pour lesquels $\varphi(\beta)$ n'est pas définie possède — s'il n'est pas vide — un premier élément $\beta_0 \in B'$, car c'est un ensemble d'ordinaux. L'application φ est alors définie sur $[0, \beta_0[$, donc en vertu de b) l'élément $\varphi(\beta)$ est aussi défini, i.e. $\beta_0 \notin B'$. Cette contradiction prouve que $B' = \varnothing$.]

Soit B^* l'ensemble, éventuellement vide, de tous les nombres β tels que $\varphi(\beta) = *$ et soit $\alpha = \sup(B \setminus B^*)$. Par construction, φ applique l'intervalle $[0, \alpha[$ sur l'ensemble A et est une bijection, i.e. φ définit une indexation de l'ensemble A par les nombres de l'intervalle $[0, \alpha[$. □

* * *

Tout ensemble A d'ordinaux est un espace topologique pour le *topologie de l'ordre* dont une base d'ensembles ouverts est constituée des intersections de l'ensemble A avec des intervalles de la forme $]\alpha, \beta[$, où $\alpha < \beta$ (par définition, chaque intervalle $]\alpha, \beta[$ est composé des nombres γ tels que $\alpha < \gamma < \beta$).

Cet espace est visiblement séparé.

Remarque 1. La topologie de l'ordre peut de toute évidence être introduite sur tout ensemble totalement ordonné; par exemple sur l'ensemble $[0, \Omega[\times [0, 1[$ muni de l'ordre *lexicographique* (i.e. d'un ordre tel que $(\alpha_1, t_1) < (\alpha_2, t_2)$ si et seulement si $\alpha_1 < \alpha_2$ ou $\alpha_1 = \alpha_2$ et $t_1 < t_2$). L'espace topologique ainsi obtenu s'appelle *longue demi-droite d'Alexandrov*. On peut se l'imaginer comme le résultat du recollement du segment $[0, 1]$ entre deux ordinaux voisins α et $\alpha + 1$ de deuxième classe.

Remarque 2. On démontre que la longue demi-droite privée de son origine $(0, 0)$ est munie d'une structure différentiable de classe C^ω qui en fait une variété séparée à une dimension ne satisfaisant pas au

deuxième axiome de dénombrabilité. (Comparer avec la remarque 5 de la leçon 7.)

La topologie de l'ordre sur les ensembles d'ordinaux possède la propriété remarquable — et un peu inattendue — suivante: *pour tout ordinal* ξ *le segment* $[0, \xi]$ *est un espace compact pour la topologie de l'ordre*. En effet, soit $\mathfrak{U}$ un recouvrement ouvert quelconque du segment $[0, \xi]$. On dira qu'un nombre $\eta \leqslant \xi$ est *admissible* si le segment $[0, \eta]$ peut être recouvert par une sous-famille finie de $\mathfrak{U}$. L'ensemble des nombres admissibles n'est pas vide (il contient 0) et s'il n'épuise pas le segment $[0, \xi]$ tout entier, il existe — en raison de la bonne ordonnance des ordinaux — un plus petit nombre non admissible $\gamma \leqslant \xi$. Soit U_0 un élément de $\mathfrak{U}$ contenant le nombre γ et soit $]\alpha, \beta[$ un intervalle tel que $\gamma \in]\alpha, \beta[\subset U_0$. Puisque $\alpha < \gamma$, le nombre α est admissible et donc le segment $[0, \alpha]$ est recouvrable par une sous-famille finie de $\mathfrak{U}$. En ajoutant U_0 à cette sous-famille, on obtient une sous-famille finie recouvrant l'intervalle $[0, \beta[= [0, \alpha] \cup]\alpha, \beta[$, donc le segment $[0, \gamma] \subset [0, \beta[$. Donc, le nombre γ est admissible contrairement à l'hypothèse. Cette contradiction prouve que tous les nombres du segment $[0, \xi]$ sont admissibles. En particulier, le nombre ξ. Donc, le recouvrement $\mathfrak{U}$ contient un sous-recouvrement fini. □

* * *

L'espace topologique $\mathscr{X} = [0, \xi]$ est tel que *chacun de ses points possède une base de voisinages qui sont des ensembles aussi bien ouverts que fermés* (à propos de ces espaces on écrit $\operatorname{ind} \mathscr{X} = 0$). Pour le prouver il suffit de remarquer que si le nombre β n'est pas limite (i.e. $\beta = \gamma + 1$), alors $]\alpha, \beta[= [\alpha + 1, \gamma]$ pour tout α. □

Bien plus, on a une proposition analogue non seulement pour les points, mais aussi pour tous les ensembles fermés, i.e. *pour chaque ensemble fermé* $F \subset [0, \xi]$ *et tout voisinage* U *de* F *il existe un voisinage* $V \subset U$ *qui est un ensemble fermé* (à propos de ces espaces $\mathscr{X}$ on écrit $\operatorname{Ind} \mathscr{X} = 0$). En effet, d'après la proposition précédente, tout point $\alpha \in F$ possède un voisinage fermé V_α contenu dans U. Les voisinages V_α forment un recouvrement ouvert de F dont on peut extraire un sous-recouvrement fini $\{V_{\alpha_1}, \ldots, V_{\alpha_n}\}$, puisque F est compact en tant que sous-ensemble fermé d'un espace compact. La réunion $V = V_{\alpha_1} \cup \ldots \cup V_{\alpha_n}$ des éléments de ce sous-recouvrement sera un voisinage fermé de F contenu dans U. □

En fait, nous avons prouvé que *si un espace topologique* $\mathscr{X}$ *est compact et* $\operatorname{ind} \mathscr{X} = 0$, *alors* $\operatorname{Ind} \mathscr{X} = 0$. [A propos, *si un espace* $\mathscr{X}$ *est séparé et* $\operatorname{Ind} \mathscr{X} = 0$, *alors* $\operatorname{ind} \mathscr{X} = 0$. Pour le prouver, il suffit de remarquer que dans un espace séparé tout singleton est fermé.]

D'autre part, il est aisé de voir que *si* $\operatorname{Ind} \mathscr{X} = 0$, *alors* $\dim \mathscr{X} = 0$. En effet, l'égalité $\dim \mathscr{X} = 0$ exprime que dans tout recouvre-

ment ouvert fini $\{U_1, \ldots, U_n\}$ de l'espace $\mathscr{X}$ on peut inscrire un recouvrement ouvert fini $\{V_1, \ldots, V_m\}$ qui est composé d'ensembles ouverts disjoints et donc (puisque le complémentaire de tout élément V_j du recouvrement $\{V_1, \ldots, V_m\}$ est la réunion des autres éléments de ce recouvrement et donc est ouvert) qui est tel que chacun de ses éléments V_j est ouvert et fermé.

Nous allons prouver plus (pour l'espace $\mathscr{X}$ avec $\operatorname{Ind} \mathscr{X} = 0$), savoir que *le nombre m des éléments du recouvrement* $\{V_1, \ldots, V_m\}$ *peut être considéré égal au nombre n des éléments du recouvrement* $\{U_1, \ldots, U_n\}$ *et de plus ce recouvrement peut être construit de telle sorte que l'on ait l'inclusion* $V_i \subset U_i$ *pour tout* $i = 1, \ldots, n$.

Pour $n = 1$ on posera $V_1 = U_1$.

Supposons que le recouvrement $\{V_1, \ldots, V_n\}$ a été construit pour tout recouvrement $\{U_1, \ldots, U_n\}$ composé de $n \geqslant 1$ ensembles ouverts. Considérons un recouvrement ouvert quelconque $\{U_1, \ldots, U_n, U_{n+1}\}$ composé de $n + 1$ ensembles. D'après l'hypothèse de la récurrence, appliquée au recouvrement $\{U_1, \ldots, U_{n-1}, U_n \cup U_{n+1}\}$, il existe un recouvrement ouvert $\{V_1, \ldots, V_{n-1}, V'_n\}$ composé de n ensembles disjoints tel que

$$V_1 \subset U_1, \ \ldots, \ V_{n-1} \subset U_{n-1}, \quad V'_n \subset U_n \cup U_{n+1}.$$

D'autre part, l'ensemble

$$F = \mathscr{X} \setminus (U_1 \cup \ldots \cup U_n)$$

étant fermé et contenu dans U_{n+1}, il existe en vertu de la condition $\operatorname{Ind} \mathscr{X} = 0$ un ensemble V à la fois ouvert et fermé tel que $F \subset V \subset U_{n+1}$. Posons

$$V_n = V'_n \cap (\mathscr{X} \setminus V), \quad V_{n+1} = V'_n \cap V.$$

Pour achever la démonstration il reste à remarquer que tous les ensembles $V_1, \ldots, V_{n-1}, V_n, V_{n+1}$ sont ouverts, disjoints, recouvrent $\mathscr{X}$ et sont tels que $V_i \subset U_i$ pour tout $i = 1, \ldots, n + 1$. □

On voit en particulier que *pour tout ordinal* ξ

$$\dim [0, \xi] = 0.$$

Il s'avère que non seulement $\operatorname{Ind} \mathscr{X} = 0$ entraîne $\dim \mathscr{X} = 0$, mais réciproquement $\dim \mathscr{X} = 0$ implique $\operatorname{Ind} \mathscr{X} = 0$ (de sorte que $\dim \mathscr{X} = 0$ *si et seulement si* $\operatorname{Ind} \mathscr{X} = 0$). En effet, si $\dim \mathscr{X} = 0$ et si U est un voisinage d'un ensemble fermé $F \subset \mathscr{X}$, alors, puisque les ensembles U et $\mathscr{X} \setminus F$ forment un recouvrement ouvert fini de $\mathscr{X}$, on peut inscrire dans ce recouvrement un recouvrement fini $\{V_j\}$ composé d'ensembles disjoints ouverts, donc fermés. La réunion V de tous les éléments de ce recouvrement rencontrant F (donc contenus dans U) sera alors un voisinage à la fois ouvert et fermé de F, contenu dans U. □

On démontre sans peine maintenant que *si des espaces* $\mathscr{X}$ *et* $\mathscr{Y}$ *sont séparés et compacts, et* dim $\mathscr{X} = 0$ *et* dim $\mathscr{Y} = 0$, *alors*

$$\dim\ (\mathscr{X} \times \mathscr{Y}) = 0$$

(de sorte que ces espaces sont justiciables de la formule (8) de la leçon 8). En effet, d'après les remarques faites ci-dessus sur les liens entre ind, Ind et dim, pour les espaces $\mathscr{X}$ et $\mathscr{Y}$ on a les égalités ind $\mathscr{X} = 0$ et ind $\mathscr{Y} = 0$, et pour l'espace $\mathscr{X} \times \mathscr{Y}$ l'égalité $\dim(\mathscr{X} \times \mathscr{Y}) = 0$ équivaut à $\operatorname{ind}(\mathscr{X} \times \mathscr{Y}) = 0$. D'autre part, il est clair que *si* ind $\mathscr{X} = 0$ *et* ind $\mathscr{Y} = 0$, *alors* $\operatorname{ind}(\mathscr{X} \times \mathscr{Y}) = 0$ (puisque, si U est un voisinage fermé d'un point p de $\mathscr{X}$, et V, un voisinage fermé d'un point q de $\mathscr{Y}$, le produit $U \times V$ sera un voisinage fermé du point (p, q) dans $\mathscr{X} \times \mathscr{Y}$ car il possède un complémentaire ouvert $(\mathscr{X} \setminus U) \times \mathscr{Y} \cup \mathscr{X} \times (\mathscr{Y} \setminus V)$.□

Remarque 3. Les notations nous incitent à penser que parallèlement à l'invariant dim (et à l'invariant « amélioré » dim′) on peut définir d'autres invariants entiers ind et Ind susceptibles aussi de faire office de dimensions. C'est effectivement le cas, mais malheureusement ces invariants sont généralement confondus pour les seuls espaces métrisables satisfaisant au deuxième axiome de dénombrabilité.

L'étude de ces invariants (et d'autres analogues) fait l'objet de la théorie des dimensions.

* * *

Après ce travail préparatoire on peut passer directement à la construction de l'exemple de Tykhonov.

Soit $\mathscr{X} = [0, \omega] \times [0, \Omega]$ le produit direct des segments d'ordinaux $[0, \omega]$ et $[0, \Omega]$. En vertu des résultats acquis ci-dessus, l'espace $\mathscr{X}$ est un espace séparé et compact de dimension 0. Soit A un sous-espace ouvert de $\mathscr{X}$ composé de tous les points de $\mathscr{X}$, sauf du point (ω, Ω). Nous allons montrer que $\dim A \geqslant 1$. Puisque $\dim A \geqslant 1$ exprime que $\dim A \neq 0$ et que $\dim A = 0$ équivaut à $\operatorname{Ind} A = 0$, il nous suffit de montrer que $\operatorname{Ind} A \neq 0$, i.e. que dans le sous-espace A il existe un ensemble fermé F et un voisinage U de F, tels qu'aucun voisinage V de F contenu dans U n'est un ensemble fermé. Nous prouverons même plus, plus exactement qu'il existe dans A un ensemble fermé F et un voisinage U de F, tels que *pour tout voisinage V de A, contenu dans U, l'adhérence $\bar{V}$ n'est pas contenue dans U.* (Ceci exprime qu'il n'existe pas de cloison entre F et $A \setminus U$ dans A, i.e. A n'est pas normal. D'autre part, l'espace $\mathscr{X}$ est normal, car compact et séparé. Donc, *l'espace $\mathscr{X}$ est aussi un exemple d'espace normal mais non complètement normal.*)

Pour F nous prenons l'ensemble de tous les points de l'espace $\mathscr{X}$ de la forme (n, Ω), où $0 \leqslant n < \omega$, et pour U, celui de tous les points

de la forme (n, β), où $0 \leqslant n < \omega$ et $0 \leqslant \beta \leqslant \Omega$. Il est clair que l'ensemble U est ouvert (dans A et dans $\mathscr{X}$) et l'ensemble F, fermé (dans A, mais pas dans $\mathscr{X}$). Soit V un ensemble ouvert (dans A, donc dans $\mathscr{X}$) contenant F et contenu dans U. Pour tout n, $0 \leqslant n < \omega$, l'ensemble des nombres $\beta \leqslant \Omega$ tels que $(n, \beta) \in V$ est ouvert dans $[0, \Omega]$ (pourquoi?) et contient Ω. Il existe donc un nombre $\alpha_n < \Omega$ tel que l'intervalle $]\alpha_n, \Omega] =]\alpha_n, \Omega[\cup \{\Omega\}$ est entièrement compris dans cet ensemble. Soit $\alpha = \sup \alpha_n$. Puisque $\alpha_n < \Omega$, on a $\alpha < \Omega$ (cf. plus haut) et $(n, \beta) \in V$ pour tout $\beta \in]\alpha, \Omega[$.

Par ailleurs, chaque voisinage W du point (ω, β) dans A contient un voisinage de la forme $[N, \omega] \times]\beta_1, \beta_2[$, où N est un entier naturel ($< \omega$) et β_1, β_2 des ordinaux de deuxième classe tels que $\beta_1 < \beta < \beta_2$. Le voisinage W contient donc tous les points (n, β) tels que $n > N$ et rencontre par conséquent V. Donc, $(n, \omega) \in \bar{V}$ et, par suite, $\bar{V} \not\subset U$. □

* * *

Revenons maintenant aux produits directs et généralisons leur construction au cas d'un nombre quelconque (généralement infini) de facteurs.

Soit $\{\mathscr{X}_\alpha, \alpha \in A\}$ une famille quelconque d'espaces topologiques et soit $\mathscr{X}$ l'ensemble de toutes les familles $\{x_\alpha\}$, où $x_\alpha \in \mathscr{X}_\alpha$.

Pour tout ensemble f i n i d'indices $\{\alpha_1, \ldots, \alpha_n\}$ et tous sous-ensembles ouverts $U_{\alpha_1} \subset \mathscr{X}_{\alpha_1}, \ldots, U_{\alpha_n} \subset \mathscr{X}_{\alpha_n}$, désignons par $O(U_{\alpha_1}, \ldots, U_{\alpha_n})$ le sous-ensemble de $\mathscr{X}$ composé de tous les points $\{x_\alpha\}$ tels que

$$x_{\alpha_1} \in U_{\alpha_1}, \ \ldots, \ x_{\alpha_n} \in U_{\alpha_n}.$$

Il est clair que

$$O(U_{\alpha_1}, \ldots, U_{\alpha_n}) = O(U_{\alpha_1}) \cap \ldots \cap O(U_{\alpha_n}),$$

donc l'intersection de deux ensembles quelconques de la forme $O(U_{\alpha_1}, \ldots, U_{\alpha_n})$ est aussi un ensemble de cette forme. Donc, l'ensemble de tous les ensembles $O(U_{\alpha_1}, \ldots, U_{\alpha_n})$ est une base d'ouverts pour une topologie sur $\mathscr{X}$ (les ouverts de cette topologie sont donc toutes les réunions des ensembles $O(U_{\alpha_1}, \ldots, U_{\alpha_n})$).

Définition 1. L'espace topologique $\mathscr{X}$ obtenu s'appelle *produit direct* (ou *tykhonovien*) des espaces $\mathscr{X}_\alpha$. On le désigne par le symbole

$$\prod_{\alpha \in A} \mathscr{X}_\alpha$$

(on se sert aussi de la notation $\mathop{\times}\limits_{\alpha \in A} \mathscr{X}_\alpha$). Si la famille $\{\mathscr{X}_\alpha\}$ est finie, on obtient de toute évidence un produit direct au sens de la définition 5 de la leçon 8.

Théorème 1. *Le produit direct de toute famille d'espaces compacts est compact.*

Ce théorème est connu sous le nom de théorème de Tykhonov. Pour le prouver, on redéfinira la compacité d'un espace dans des termes qui rendront la démonstration plus commode.

* * *

Supposons pour l'instant que $\mathscr{X}$ est un ensemble arbitraire.

Définition 2. On dit qu'un ensemble Φ de parties de $\mathscr{X}$ est un *filtre* si

1° toute partie de $\mathscr{X}$ contenant un élément de Φ appartient à Φ:

$$A \subset B \subset \mathscr{X}, \quad A \in \Phi \Rightarrow B \in \Phi;$$

2° l'intersection de deux éléments quelconques de Φ appartient à Φ:

$$A, B \in \Phi \Rightarrow A \cap B \in \Phi;$$

3° $\mathscr{X}$ appartient à Φ mais pas l'ensemble $\varnothing$;

$$\mathscr{X} \in \Phi, \quad \varnothing \notin \Phi.$$

Exemple 1. Pour tout point x_0 de $\mathscr{X}$ la famille de toutes les parties de $\mathscr{X}$ contenant x_0 est un filtre.

Exemple 2. Si $\mathscr{X}$ est un espace topologique, pour tout point $x_0 \in \mathscr{X}$ la famille de toutes les parties A de $\mathscr{X}$ pour lesquelles x_0 est un point intérieur ($x_0 \in \operatorname{Int} A$) est un filtre.

Exemple 3. Si l'ensemble est infini, tous les ensembles dont les complémentaires sont finis forment un filtre.

Exemple 4. Si Φ est un filtre et M un sous-ensemble n'appartenant pas à Φ, l'ensemble Φ_M de toutes les parties A de $\mathscr{X}$ telles que $A \cup M \in \Phi$ est un filtre.

* * *

Définition 3. On dit qu'un ensemble Γ de parties d'un ensemble $\mathscr{X}$ est *centré* si l'intersection de toute sous-famille finie d'éléments de Γ n'est pas vide.

Il est clair que *si Γ est centré, il en est de même de l'ensemble $f\Gamma = \{fA\,;\ A \in \Gamma\}$ pour toute application $f\colon \mathscr{X} \to \mathscr{Y}$.*

Des propriétés 2 et 3 des filtres il s'ensuit immédiatement que *tout filtre est un ensemble centré.*

Réciproquement, il est aisé de voir que *tout ensemble centré Γ est contenu dans un filtre.* En effet, il est clair qu'un tel filtre sera par exemple l'ensemble $\Phi(\Gamma)$ de toutes les parties de $\mathscr{X}$ qui contiennent chacune un sous-ensemble qui est l'intersection d'une famille finie d'éléments de l'ensemble Γ. □

Le filtre Φ (Γ) est de toute évidence le filtre minimal contenant l'ensemble centré Γ. On dit qu'il est *engendré* par cet ensemble.

* * *

On rappelle qu'un élément m d'un ensemble ordonné A est *maximal* si $m \leqslant a$ entraîne $m = a$. Deux éléments a et b d'un ensemble A sont *comparables* si ou bien $a \leqslant b$, ou bien $b \leqslant a$. Un sous-ensemble C d'un ensemble ordonné est une *chaîne* s'il est totalement ordonné, i.e. si deux quelconques de ses éléments sont comparables. Un élément c_0 est un *majorant* d'une chaîne C si $c \leqslant c_0$ pour tout $c \in C$. Un ensemble ordonné A est *inductif* si toute chaîne de A admet au moins un majorant.

Un exemple d'ensemble ordonné est l'ensemble FILTR($\mathscr{X}$) de tous les filtres d'un ensemble $\mathscr{X}$, ordonné par l'inclusion ($\Phi \leqslant \Psi$ si $\Phi \subset \Psi$).

Les éléments maximaux de cet ensemble s'appellent *ultrafiltres*. Donc, un filtre Φ est un ultrafiltre si tout filtre qui le contient est confondu avec Φ.

Il est aisé de voir qu'*un ensemble Φ de parties d'un ensemble $\mathscr{X}$ est un ultrafiltre si et seulement si*

a) *l'intersection de deux ensembles quelconques de Φ n'est pas vide*;

b) *l'ensemble $\mathscr{X}$ appartient à Φ*;

c) *l'un de deux sous-ensembles complémentaires quelconques A et $B = \mathscr{X} \setminus A$ de l'ensemble $\mathscr{X}$ (et lui seul) appartient à Φ.*

En effet, la condition a) est la condition 2° de la définition d'un filtre, la condition b) est la première partie de la condition 3° dont la deuxième partie résulte de la première et de la condition c) (appliquée à $A = \mathscr{X}$). Si $A \subset B \subset \mathscr{X}$ et $A \in \Phi$, mais $B \notin \Phi$, la condition c) nous donne $\mathscr{X} \setminus B \in \Phi$, ce qui contredit la condition a), puisque $(\mathscr{X} \setminus B) \cap A = \varnothing$. Donc, l'ensemble Φ qui vérifie les conditions a), b) et c) est un filtre. Si ce filtre est contenu dans un filtre Ψ et s'il existe un ensemble $A \in \Psi$, n'appartenant pas à Φ, alors $B = \mathscr{X} \setminus A$ appartient à Φ, donc à Ψ aussi, ce qui est impossible (car $A \cap B = \varnothing$). Donc, $\Psi = \Phi$, et, par suite, Φ est un ultrafiltre.

Réciproquement, si Φ est un ultrafiltre, les conditions a) et b) sont visiblement satisfaites et, pour prouver la condition c), il suffit de remarquer que pour $A \notin \Phi$ le filtre Φ_A (cf. exemple 4) contient de toute évidence le filtre Φ. Donc, puisque l'ultrafiltre Φ est maximal, le filtre Φ_A est confondu avec Φ. Mais il est clair que le filtre Φ_A contient l'ensemble $B = \mathscr{X} \setminus A$. Donc, $B \in \Phi$. □

En particulier, il s'ensuit immédiatement que *chaque filtre de l'exemple* 1 *est un ultrafiltre*. Ces ultrafiltres sont dits *triviaux*.

Il est aisé de voir que *tout ensemble A rencontrant chaque élément de l'ultrafiltre Φ appartient à Φ*. En effet, si $A \notin \Phi$, il existe dans

l'ultrafiltre Φ un élément (plus exactement $\mathscr{X} \setminus A$) ne rencontrant pas A. $\square$

Les chaînes de l'ensemble FILTR ($\mathscr{X}$) sont des ensembles de filtres $\{\Phi_\alpha\}$ tels que pour tous α et β ou bien $\Phi_\alpha \subset \Phi_\beta$, ou bien $\Phi_\beta \subset \Phi_\alpha$. Ceci étant, il est aisé de voir que *la réunion Φ de tous les filtres d'une chaîne est un filtre* (il suffit de remarquer que deux éléments quelconques de cette réunion appartiennent à un filtre Φ_α). Puisque le filtre Φ contient tous les filtres Φ_α, il en est un majorant. On voit donc que *l'ensemble de tous les filtres est inductif.*

Mais d'après le lemme classique de Zorn (appelé aussi théorème de Kuratowski-Zorn), *tout ensemble inductif A admet un élément maximal et, bien plus, pour tout élément $a_0 \in A$ il existe un élément maximal m tel que $a_0 \leqslant m$.*

Prouvons ce lemme. Supposons que l'ensemble A est arbitrairement numéroté par les ordinaux d'un intervalle $[0, \alpha[$. (Signalons que cette numérotation n'est en aucune façon liée à l'ordre de A.) Il est clair que sans perdre en généralité on peut admettre que l'élément a_0 possède le numéro 0. Considérons dans A un sous-ensemble C satisfaisant aux conditions suivantes:

a) l'élément a_0 appartient à C;

b) l'élément de A de numéro ξ appartient à C si et seulement s'il est comparable à tous les éléments de A de numéro inférieur.

Il est clair que ces conditions définissent de façon unique le sous-ensemble C et que C est une chaîne. Soit m un majorant de C (ce majorant existe puisque A est inductif). L'élément m étant comparable à tout élément de C, tout élément $a \geqslant m$ est comparable à tout élément de C et, en particulier, à tout élément de C de numéro inférieur. Donc, d'après la condition b), l'élément a appartient à C et, par suite, $a \leqslant m$. Ceci prouve que m est maximal. Pour achever la démonstration, il reste à remarquer que $a_0 \leqslant m$ par construction.

D'après le lemme de Zorn appliqué à l'ensemble inductif FILTR ($\mathscr{X}$), *tout filtre de $\mathscr{X}$ est contenu dans un ultrafiltre.*

Ceci nous fournit une source intarissable d'ultrafiltres non triviaux.

Remarque 4. Il est curieux de noter qu'aucun ultrafiltre non trivial ne peut être explicitement construit sans l'introduction d'éléments arbitraires, i.e., en d'autres termes, ne peut être caractérisé par des propriétés qui le définissent de façon unique. (En logique mathématique, on démontre cette proposition en définissant rigoureusement toutes les notions logico-mathématiques nécessaires.)

* * *

Supposons maintenant que $\mathscr{X}$ est un espace topologique.

Définition 4. On dit qu'un point $p_0 \in \mathscr{X}$ est un *point adhérent* d'un ensemble Φ de parties d'un espace topologique $\mathscr{X}$ si chacun de ses voisinages rencontre chaque élément A de Φ, i.e., en d'autres termes, si $p_0 \in \bar{A}$ pour tout $A \in \Phi$.

Donc, les points adhérents de l'ensemble Φ sont les points de l'intersection

$$\bigcap_{A \in \Phi} \bar{A} \tag{2}$$

des adhérences $\bar{A}$ de tous les ensembles $A \in \Phi$.

Remarquons que *si Φ est un filtre, l'ensemble* (2) *est confondu avec l'intersection de tous les ensembles fermés de Φ.*

Le complémentaire de l'ensemble (2) est la réunion des ensembles ouverts $\mathscr{X} \setminus \bar{A}$. Donc, *l'ensemble Φ ne possède pas de points adhérents si et seulement si les ensembles $\mathscr{X} \setminus \bar{A}$, $A \in \Phi$, forment un recouvrement de $\mathscr{X}$.*

Définition 5. Si tout voisinage d'un point $p_0 \in \mathscr{X}$ contient un élément de l'ensemble des parties de Φ (et donc appartient à Φ si Φ est un filtre), on dit que Φ *converge vers le point p_0*.

Il est aisé de voir que *si l'ensemble Φ est un ultrafiltre, il converge vers chacun de ses points adhérents.* En effet, si un voisinage U de p_0 rencontre chaque élément de l'ultrafiltre Φ, alors U appartient à Φ. □

Nous pouvons prouver maintenant le critère de compacité.

Proposition 1. *Les propriétés suivantes de l'espace topologique $\mathscr{X}$ sont équivalentes*:

1° *L'espace $\mathscr{X}$ est compact.*

2° *Tout ensemble centré de parties fermées de $\mathscr{X}$ possède une intersection non vide.*

3° *Tout filtre de $\mathscr{X}$ possède des points adhérents.*

4° *Tout ultrafiltre de $\mathscr{X}$ est convergent.*

D é m o n s t r a t i o n. Il suffit de toute évidence de prouver les implications

$$\begin{tikzcd}
 & & 3^\circ \arrow[d, Rightarrow] \\
1^\circ \arrow[r, Leftrightarrow] & 2^\circ \arrow[ur, Rightarrow] & 4^\circ \arrow[l, Rightarrow]
\end{tikzcd}$$

I m p l i c a t i o n s $1^\circ \Rightarrow 2^\circ$ et $2^\circ \Rightarrow 1^\circ$. Dire que l'intersection de tous les éléments A d'un ensemble Γ de parties de l'espace $\mathscr{X}$ est vide revient à dire que les complémentaires $\mathscr{X} \setminus A$ de ces éléments forment un recouvrement de $\mathscr{X}$. Quant à l'affirmation que l'ensemble Γ est centré, elle exprime qu'aucune famille finie de complémentaires $\mathscr{X} \setminus A$ ne forme un recouvrement de $\mathscr{X}$. Donc, si un ensemble de parties fermées d'un espace compact $\mathscr{X}$ possède une intersection vide, il ne peut en aucun cas être centré, et si l'intersection de tout

ensemble centré de parties fermées de l'espace $\mathcal{X}$ n'est pas vide, chaque famille d'ensembles ouverts de $\mathcal{X}$, dont aucune sous-famille finie n'est un recouvrement de $\mathcal{X}$, n'est pas elle-même un recouvrement de $\mathcal{X}$.

Implication $2^\circ \Rightarrow 3^\circ$. Il suffit de remarquer que l'ensemble Γ de toutes les parties fermées appartenant à un filtre Φ est centré et que chaque point adhérent du filtre Φ est un point commun aux éléments de l'ensemble Γ.

Implication $3^\circ \Rightarrow 4^\circ$. Si chaque filtre possède des points adhérents, il en est de même en particulier de chaque ultrafiltre. D'autre part, un ultrafiltre converge si et seulement s'il possède des points adhérents.

Implication $4^\circ \Rightarrow 2^\circ$. Chaque ensemble centré est contenu dans un ultrafiltre. Donc, si cet ultrafiltre converge, l'ensemble centré possède un point adhérent (tout point adhérent d'un ensemble quelconque de parties est un point adhérent de chaque partie de cet ensemble). D'autre part, si un ensemble de parties est composé d'ensembles fermés, chacun de ses points adhérents appartient à l'intersection de tous ses éléments. Donc, si chaque ultrafiltre de l'espace $\mathcal{X}$ converge (possède un point adhérent), tout ensemble centré de parties fermées de $\mathcal{X}$ possède une intersection non vide. □

* * *

On peut maintenant établir le théorème de Tykhonov sans aucune peine.

Démonstration du théorème 1. Soit Φ un ultrafiltre sur un produit direct

$$\mathcal{X} = \prod_{\alpha \in A} \mathcal{X}_\alpha$$

d'espaces compacts $\mathcal{X}_\alpha$. Considérons l'image $\mathrm{pr}_\alpha \Phi$ de cet ultrafiltre par la projection canonique $\mathrm{pr}_\alpha : \mathcal{X} \to \mathcal{X}_\alpha$, $\{x_\alpha\} \mapsto x_\alpha$. L'ensemble Φ étant centré, il en est de même de sa projection $\mathrm{pr}_\alpha \Phi$ donc aussi de l'ensemble Γ_α des adhérences des éléments de l'ensemble $\mathrm{pr}_\alpha \Phi$ (i.e. de $\overline{A}_\alpha$, où $A_\alpha = \mathrm{pr}_\alpha A$, $A \in \Phi$). Par conséquent, puisque l'espace $\mathcal{X}_\alpha$ est compact, il existe un point $p_\alpha \in \mathcal{X}_\alpha$ appartenant à tous les ensembles $\overline{A}_\alpha$ de Γ_α, i.e. un point tel que chacun de ses voisinages U_α rencontre chaque ensemble A_α.

Pour les ensembles A du filtre Φ, ceci exprime qu'ils rencontrent tous l'ensemble $O(U_\alpha)$, d'où il s'ensuit, puisque Φ est un ultrafiltre, que $O(U_\alpha) \in \Phi$.

Considérons un voisinage W du point $\{p_\alpha\} \in \mathcal{X}$ obtenu en prenant un point p_α pour chaque α. Par définition, ce voisinage contient un ensemble de la forme $O(U_{\alpha_1}, \ldots, U_{\alpha_n})$ où $U_{\alpha_1}, \ldots, U_{\alpha_n}$

sont des voisinages des points $p_{\alpha_1}, \ldots, p_{\alpha_n}$ dans les espaces respectifs $\mathcal{X}_{\alpha_1}, \ldots, \mathcal{X}_{\alpha_n}$. Mais nous venons tout juste de voir que les ensembles $O(U_{\alpha_1}), \ldots, O(U_{\alpha_n})$ appartiennent à l'ultrafiltre Φ. Donc, Φ contient aussi leur intersection

$$O(U_{\alpha_1}, \ldots, U_{\alpha_n}) = O(U_{\alpha_1}) \cap \ldots \cap O(U_{\alpha_n}).$$

et, par conséquent, le voisinage W.

Donc, tout voisinage de $\{p_\alpha\}$ appartient à Φ et, par suite, Φ converge vers $\{p_\alpha\}$.

Ceci prouve que $\mathcal{X}$ est compact, car tout ultrafiltre y converge. $\square$

LEÇON 11

Structure différentiable sur un espace affine. — Variété de matrices de rang donné. — Variétés de Stiefel. — Séries de matrices. — Exponentielle d'une matrice. — Logarithme d'une matrice. — Matrices orthogonales et J-orthogonales. — Groupes de Lie de matrices. — Groupes de matrices J-orthogonales. — Matrices unitaires et J-unitaires. — Groupes de Lie de matrices complexes. — Variétés $\mathbb{C}$-analytiques. — Espaces connexes par arcs. — Espaces connexes. — Coïncidence de la connexité et de la connexité par arcs pour les variétés. — Chemins différentiables et différentiables par morceaux. — Variétés connexes ne satisfaisant pas au deuxième axiome de dénombrabilité.

Revenons aux variétés différentiables (de classe C^∞) et complétons la liste des exemples citée à la leçon 6. Non seulement ces exemples illustrent la théorie générale, mais ils sont la source d'importants théorèmes qui malheureusement sortent presque entièrement du cadre de ce cours.

* * *

Soit $\mathcal{A}$ un espace affine à n dimensions sur le corps $\mathbb{R}$. Tout repère $O\mathbf{e}_1 \ldots \mathbf{e}_n$ de cet espace définit un isomorphisme de coordonnées $h: \mathcal{A} \to \mathbb{R}^n$ (associant à un point M, $\overrightarrow{OM} = x^1\mathbf{e}_1 + \ldots + x^n\mathbf{e}_n$, le point $(x^1, \ldots, x^n)$ de $\mathbb{R}^n$), donc une carte $(\mathcal{A}, h)$.

Cette carte forme un atlas à un seul élément sur $\mathcal{A}$ et, par conséquent, définit une structure différentiable sur $\mathcal{A}$ (comparer avec l'exemple 1 de la leçon 6). Pour tout autre repère $O'\mathbf{e}'_1 \ldots \mathbf{e}'_n$, la carte correspondante $(\mathcal{A}, h')$ est liée à la carte $(\mathcal{A}, h)$ par l'application de passage $h' \circ h^{-1}$ qui est une application affine (donc un difféomorphisme de classe C^ω) de $\mathbb{R}^n$ sur $\mathbb{R}^n$ et, par conséquent, est compatible avec la carte $(\mathcal{A}, h)$. Ceci exprime que la structure différentiable définie par la carte $(\mathcal{A}, h)$ ne dépend pas du choix du repère $O\mathbf{e}_1 \ldots \mathbf{e}_n$. Nous l'appellerons *structure différentiable standard* sur $\mathcal{A}$. Elle est manifestement compatible avec la topologie de $\mathcal{A}$.

Puisque tout espace vectoriel est automatiquement affine, tout ce qui vient d'être dit est valable en particulier pour un espace vectoriel quelconque, par exemple pour l'espace vectoriel à mn dimensions $\mathbb{R}(m, n)$ de toutes les $(m \times n)$-matrices rectangulaires réelles. Donc, $\mathbb{R}(m, n)$ *est une variété différentiable (de classe C^ω) à mn dimensions.* Pour obtenir une carte recouvrant cette variété, il faut écrire chaque matrice A de $\mathbb{R}(m, n)$ sous forme d'une ligne dans un ordre quelconque qui soit le même pour toutes les matrices.

Le groupe GL (n) des matrices régulières d'ordre n est caractérisé dans l'espace vectoriel $\mathrm{Mat}_n(\mathbb{R}) = \mathbb{R}(n, n)$ des matrices carrées A d'ordre n par la condition $\det A \neq 0$. Puisque le déterminant d'une matrice dépend continûment des éléments de cette matrice (la fonction $\det : A \to \det A$ est continue sur $\mathrm{Mat}_n(\mathbb{R})$), le groupe GL (n) est un ensemble ouvert de la variété différentiable $\mathrm{Mat}_n(\mathbb{R})$ et donc est muni d'une structure différentiable induite. Par conséquent, *le groupe* GL (n) *est une variété différentiable* (*de classe* C^∞). La dimension de cette variété est n^2.

* * *

Pour obtenir un exemple moins trivial de variété différentiable, considérons le sous-ensemble $\mathbb{R}(m, n; k)$ de l'espace $\mathbb{R}(m, n)$, composé des matrices de rang k, où $0 \leqslant k \leqslant \min(m, n)$. Etant un sous-ensemble de l'espace topologique $\mathbb{R}(m, n)$, cet ensemble est un espace topologique (pour la topologie induite).

Il s'avère que *l'ensemble* $\mathbb{R}(m, n; k)$ *peut être muni de façon naturelle d'une structure de variété différentiable à* $k(m + n - k)$ *dimensions, compatible avec la topologie.*

Considérons à cet effet le sous-ensemble $U \subset \mathbb{R}(m, n; k)$ des $(m \times n)$-matrices de la forme

$$\begin{Vmatrix} A & B \\ C & D \end{Vmatrix}, \tag{1}$$

où A est une $(k \times k)$-matrice régulière (i.e. appartenant à GL (k)), et B, C et D respectivement des $(m - k) \times k$, $k \times (n - k)$ et $(m - k) \times (n - k)$-matrices arbitraires.

En écrivant les matrices A, B et C sous forme d'une ligne dans le même ordre, on obtient une application h de l'ensemble U sur un sous-ensemble U_0 de l'espace $\mathbb{R}^N$, où

$$N = k^2 + (m - k)k + k(n - k) = k(m + n - k).$$

La matrice (1) étant de même rang que la matrice

$$\begin{Vmatrix} E & 0 \\ -CA^{-1} & E \end{Vmatrix} \begin{Vmatrix} A & B \\ C & D \end{Vmatrix} = \begin{Vmatrix} A & B \\ 0 & D - CA^{-1}B \end{Vmatrix}$$

(car le premier facteur du premier membre est une matrice régulière) laquelle est de rang k si et seulement si $D - CA^{-1}B = 0$, la matrice (1) est définie de façon unique par les matrices A, B, C, où A est une matrice régulière, et B et C des matrices généralement arbitraires. Ceci exprime que l'ensemble U, donc l'ensemble U_0, sont ouverts (dans $\mathbb{R}(m, n; k)$ et $\mathbb{R}^N$ respectivement), et l'application $h : U \to U_0$ est bijective. Donc, le couple (U, h) est une carte sur $\mathbb{R}(m, n; k)$.

L'application réciproque $h^{-1}: U_0 \to U$ consiste à restituer les matrices A, B et C à l'aide de la matrice ligne donnée et à écrire dans le coin inférieur droit de la matrice (1) la matrice

$$D = CA^{-1}B.$$

Soit maintenant $\alpha : \mathbb{R}(m, n; k) \to \mathbb{R}(m, n; k)$ une permutation des lignes ou des colonnes (il existe en tout $m!n!$ permutations) et soit

$$U_\alpha = \alpha U \quad \text{et} \quad h_\alpha = h \circ \alpha^{-1}.$$

Le couple (U_α, h_α) sera alors une carte sur $\mathbb{R}(m, n; k)$.

Puisque toute $(m \times n)$-matrice de rang k peut être transformée par des permutations des lignes et des colonnes en une matrice de la forme (1), *les ensembles U_α forment un recouvrement de l'ensemble* $\mathbb{R}(m, n; k)$.

Supposons que $U_\alpha \cap U_\beta \neq \varnothing$. L'ensemble

$$\alpha^{-1}(U_\alpha \cap U_\beta) = U \cap (\alpha^{-1} \circ \beta) U$$

est composé de matrices (1) dont les images par la permutation $\alpha^{-1} \circ \beta$ appartiennent à U, i.e. dont le mineur d'ordre k se transforme par $\alpha^{-1} \circ \beta$ en le mineur supérieur gauche qui est non nul. Ce mineur est une fonction rationnelle (car $D = CA^{-1}B$), donc continue (dans son domaine de définition) des éléments des matrices A, B et C. Donc, l'ensemble sur lequel cette fonction est non nulle, i.e. l'ensemble

$$h(\alpha^{-1}(U_\alpha \cap U_\beta)) = h_\alpha(U_\alpha \cap U_\beta),$$

est ouvert dans $U_0 = h(U)$.

L'application

$$(2) \quad h_\beta \circ h_\alpha^{-1} = h \circ (\beta^{-1} \circ \alpha) \circ h^{-1} : h_\alpha(U_\alpha \cap U_\beta) \to h_\beta(U_\alpha \cap U_\beta)$$

envoie le point de $h_\alpha(U_\alpha \cap U_\beta) \subset U_0$ correspondant à une matrice (1) (plus exactement aux matrices A, B et C) en une matrice (1), i.e. en la matrice

$$\begin{Vmatrix} A & B \\ C & CA^{-1}B \end{Vmatrix},$$

puis permute les lignes et colonnes de cette matrice et enfin inscrit sur une même ligne tous les éléments se trouvant à la place des matrices A, B et C. Ceci montre que l'application (2) est définie par des fonctions rationnelles des coordonnées. Ces fonctions sont définies sur $h_\alpha(U_\alpha \cap U_\beta)$, donc sont différentiables sur $h_\alpha(U_\alpha \cap U_\beta)$. Ceci prouve que *deux cartes quelconques de la forme (U_α, h_α) sont compatibles.*

Ces cartes forment donc un atlas. La structure différentiable définie par cet atlas (structure qui est manifestement compatible avec la topologie de $\mathbb{R}(m, n; k)$) n'est autre que la structure que nous voulions construire. □

Remarquons que pour $m = n = k$ l'atlas est constitué d'une seule carte (U, h), et cet atlas définit la structure différentiable sur $\mathrm{GL}(n)$ indiquée plus haut.

* * *

Soit $\mathcal{T}$ un espace vectoriel à n dimensions sur $\mathbb{R}$. Appelons *m-repères* de $\mathcal{V}$, où $0 \leqslant m \leqslant n$, des familles libres de m vecteurs de $\mathcal{T}$ et considérons l'ensemble $V(m, \mathcal{V})$ des m-repères de $\mathcal{T}$. En choisissant une base (un n-repère) dans $\mathcal{V}$, on peut associer à chaque m-repère une $(m \times n)$-matrice dont les colonnes sont composées des coordonnées des vecteurs de ce repère dans cette base. Il est clair qu'on obtient de la sorte une application bijective $V(m, \mathcal{V}) \to \mathbb{R}(m, n; m)$. En transportant la structure différentiable de $\mathbb{R}(m, n; m)$ sur $V(m, \mathcal{V})$ à l'aide de cette application, on munit $V(m, \mathcal{V})$ d'une structure de variété différentiable qui ne dépend visiblement pas du choix de la base dans $\mathcal{V}$.

Cette variété s'appelle *variété de Stiefel* de l'espace vectoriel $\mathcal{T}$. Si $\mathcal{T} = \mathbb{R}^n$, on la désigne par $V(m, n)$.

Exercice 1. Montrer que l'ensemble $G(m, \mathcal{V})$ des sous-espaces à m dimensions d'un espace vectoriel $\mathcal{V}$ à n dimensions est une variété différentiable à $m(n - m)$ dimensions. (Cette variété s'appelle *variété grassmannienne*; cf. leçon II.10.)

* * *

Nous aurons besoin de quelques notions de théorie générale des matrices pour étudier des variétés de matrices différentiables plus complexes.

L'espace vectoriel $\mathrm{Mat}_n(\mathbb{R})$ des matrices $A = \| a_i^j \|$ d'ordre n peut être muni de plusieurs normes distinctes, i.e. de fonctions $A \mapsto |A|$ à valeurs dans $\mathbb{R}$ telles que

$$|A + B| \leqslant |A| + |B|, \qquad |kA| = |k| \cdot |A|$$

et

$$|A| > 0 \quad \text{pour} \quad A \neq 0$$

pour toutes matrices A et B et tout nombre k. La plus commode sera la norme définie par la formule

$$|A| = n \cdot \max_{1 \leqslant i,\, j \leqslant n} |a_i^j|.$$

Cette norme possède l'agréable propriété suivante: *pour toutes matrices* $A = \| a_j^i \|$ *et* $B = \| b_j^i \|$ *on a l'inégalité*

$$|AB| \leqslant |A| \cdot |B|.$$

En effet, par définition de la multiplication des matrices, les éléments c_j^i de la matrice AB s'expriment à l'aide de la formule $c_j^i = a_k^i b_j^k$, i.e.

$$c_j^i = \sum_{k=1}^{n} a_k^i b_j^k, \quad i, j = 1, \ldots, n.$$

Donc,

$$|AB| = n \cdot \max_{1 \leqslant i, j \leqslant n} |c_j^i| \leqslant n \cdot \max_{1 \leqslant i, j \leqslant n} \sum_{k=1}^{n} |a_k^i| \cdot |b_j^k| \leqslant$$

$$\leqslant n \cdot \underbrace{\left(\frac{|A|}{n} \cdot \frac{|B|}{n} + \ldots + \frac{|A|}{n} \cdot \frac{|B|}{n} \right)}_{n \text{ fois}} = |A| \cdot |B|. \ \square$$

De plus, cette norme est telle que

$$|A^{\mathrm{T}}| = |A|.$$

Conformément aux définitions générales de l'analyse on dira qu'une suite $\{A(m)\}$ de matrices $A(m) = \| a_j^i(m) \|$ *converge* vers une matrice $A = \| a_j^i \|$ (et on écrira $A(m) \to A$ lorsque $m \to \infty$) si

$$| A(m) - A | \to 0$$

lorsque $m \to \infty$ ou, ce qui est visiblement équivalent, si $a_j^i(m) \to a_j^i$ pour tous $i, j = 1, \ldots, m$.

On dira qu'une série infinie de matrices

$$A_0 + A_1 + \ldots + A_m + \ldots \tag{3}$$

converge vers une matrice A s'il en est de même de la suite $\{A(m)\}$ de ses *sommes partielles*

$$A(m) = A_0 + \ldots + A_m.$$

Le classique c r i t è r e d e c o n v e r g e n c e d e C a u c h y s'étend immédiatement aux séries de matrices, i.e. *la série* (3) *converge si et seulement si pour tout* $\varepsilon > 0$ *il existe un* $M_0 > 0$ *tel que pour tout* $m > M_0$ *et tout* $p \geqslant 0$ *on ait*

$$| A_m + A_{m+1} + \ldots + A_{m+p} | < \varepsilon.$$

On dit que la série (3) est *absolument convergente* si converge la série numérique

$$| A_0 | + | A_1 | + \ldots + | A_m | + \ldots$$

Puisque

$$|A_m + A_{m+1} + \ldots + A_{m+p}| \leqslant |A_m| + |A_{m+1}| + \ldots + |A_{m+p}|,$$

le critère de convergence de Cauchy entraîne immédiatement *la convergence de toute série absolument convergente.*

* * *

Comme exemple de série absolument convergente citons

$$\sum_{m=0}^{\infty} \frac{A^m}{m!} = E + A + \frac{A^2}{2!} + \ldots + \frac{A^m}{m!} + \ldots$$

Cette série converge (absolument) pour toute matrice A. Sa somme s'appelle *exponentielle* de cette matrice et se note e^A (ou $\exp A$).

Il est clair que l'opération de passage à la matrice e^A commute à l'opération $A \mapsto A^\top$ de transposition des matrices, i.e.

$$(4) \qquad (e^A)^\top = e^{A^\top} \text{ pour toute matrice } A.$$

De plus,

$$(5) \qquad C^{-1}e^A C = e^{C^{-1}AC}$$

pour toute matrice régulière C (car $C^{-1}A^m C = (C^{-1}AC)^m$ pour tout m).

Si la matrice A est trigonale supérieure d'éléments diagonaux $a_1, \ldots, a_n$, chaque matrice A^m sera aussi trigonale supérieure et ses éléments diagonaux seront $a_1^m, \ldots, a_n^m$. Donc, la matrice e^A sera aussi trigonale supérieure et aura pour éléments diagonaux $e^{a_1}, \ldots, e^{a_n}$. Le déterminant de la matrice e^A est alors donné par la formule

$$(6) \qquad \det e^A = e^{\operatorname{Tr} A}$$

où $\operatorname{Tr} A$ est la trace de la matrice A (la somme de ses éléments diagonaux). Mais on sait du cours d'algèbre linéaire (puisque la trace est l'un des coefficients du polynôme caractéristique) que pour toute matrice régulière,

$$(7) \qquad \operatorname{Tr}(C^{-1}AC) = \operatorname{Tr} A$$

et que (cf. proposition 5 de la leçon II.15) toute matrice est de la forme $C^{-1}AC$, où A est une matrice trigonale supérieure (généralement à éléments complexes). Donc, en vertu de la formule (5), *la formule* (6) *est valable pour toute matrice* A.

On voit en particulier que *le déterminant de chaque matrice* e^A *est strictement positif*, de sorte que, en particulier, *la matrice* e^A *est régulière* (appartient au groupe $\mathrm{GL}(n)$).

L'égalité

$$(8) \qquad e^A e^B = e^{A+B}$$

n'a pas lieu pour toutes matrices A et B. Mais *elle est valable si les matrices* A *et* B *commutent*, car dans ce cas les raisonnements qui ont servi à prouver la formule (8) pour les nombres restent en vigueur. [D'après la formule du binôme, pour tout $m \geqslant 0$

$$\frac{(A+B)^m}{m!} = \sum_{k=0}^{m} \frac{A^k}{k!} \frac{B^{m-k}}{(m-k)!},$$

donc pour tout $M > 0$

$$\sum_{m=0}^{2M} \frac{(A+B)^m}{m!} = \left(\sum_{m=0}^{M} \frac{A^m}{m!}\right)\left(\sum_{m=0}^{M} \frac{B^m}{m!}\right) + R_M,$$

où R_M est la somme des termes de la forme

$$\frac{A^k}{k!} \cdot \frac{B^l}{l!}$$

étendue à tous les couples (k, l) tels que $k + l \leqslant 2M$ et l'un des nombres k ou l est $> M$. Mais puisque le nombre de ces couples est égal à $M(M+1)$, et

$$\left|\frac{A^k}{k!} \cdot \frac{B^l}{l!}\right| \leqslant \frac{|A|^k}{k!} \frac{|B|^l}{l!} \leqslant \frac{\max(|A|, |B|)^{2M}}{M!},$$

on a

$$|R_M| \leqslant M(M+1)\frac{\max(|A|,\ |B|)^{2M}}{M!}.$$

Donc, $|R_M| \to 0$ lorsque $M \to \infty$, ce qui prouve la formule (8).]

On voit en particulier que $e^A e^{-A} = e^0 = E$, i.e.

$$(e^A)^{-1} = e^{-A}. \tag{9}$$

(On remarquera que ceci prouve une fois de plus que e^A appartient à $GL(n)$.)

* * *

Par analogie à la fonction exp, on peut définir d'autres fonctions matricielles de A. Soit par exemple la série

$$(A-E) - \frac{(A-E)^2}{2} + \ldots + (-1)^{m+1}\frac{(A-E)^m}{m} + \ldots, \tag{10}$$

qui est l'analogue matriciel de la série logarithmique. Puisque la série logarithmique converge absolument pour $|z-1| < 1$, *la série* (10) *converge absolument pour* $|A-E| < 1$. Sa somme se note $\ln A$ et s'appelle *logarithme de la matrice* A.

Soulignons que $\ln A$ *n'est ainsi définie que pour* $|A-E| < 1$.

D'une façon générale, $\ln(AB) \neq \ln A + \ln B$, mais *si les matrices* A *et* B *commutent* (et les matrices $\ln A$, $\ln B$ et $\ln(AB)$ sont définies), *alors*

$$\ln(AB) = \ln A + \ln B. \tag{11}$$

La démonstration est identique à celle de la formule (6).

Si A est un nombre (une matrice d'ordre 1), on sait alors que $e^{\ln A} = A$. Donc, en substituant la série de $\ln A$ à A dans la série de e^A, en chassant les parenthèses et en réduisant les termes semblables, on obtient A (démontrer directement cette assertion), et de plus ce calcul sera licite pour tout A tel que $\ln A$ converge. Puisque ce calcul (et la justification de sa légitimité) est valable intégralement pour les matrices de tout ordre, on voit donc que

$$e^{\ln A} = A \tag{12}$$

pour toute matrice A *telle que* $|A-E| < 1$. (En particulier, pour une telle matrice on a $\det A > 0$.)

Pour des raisons analogues on doit avoir

$$\ln e^A = A, \tag{13}$$

et de plus, puisque pour les nombres le calcul effectué sur les séries n'est licite que pour $|z| < \ln 2$ (pourquoi ?), *l'égalité* (13) *est vérifiée pour* $|A| < \ln 2$. (On remarquera que si $|A| < \ln 2$, alors $|e^A - E| \leqslant e^{|A|} - 1 < 1$.)

Des formules (12) et (13) il résulte que *les fonctions matricielles* exp *et* ln *définissent des applications bijectives réciproques l'une de l'autre d'un voisinage de la matrice unité, composé de matrices* A *de la forme* e^B, *où* $|B| < \ln 2$, *sur un voisinage de la matrice nulle* 0, *composé de matrices* B *telles que* $|B| < \ln 2$.

Ceci étant, *ces applications sont des difféomorphismes*, puisque les éléments des matrices e^A et $\ln A$ sont visiblement des fonctions différentiables ($\mathbb{R}$-analytiques) des éléments de la matrice A.

Exercice 2. Déduire la formule (11) de la formule (8). [N o t a. Montrer que pour $|A| < \ln 2$ et $|B| < \ln 2$ la commutativité des matrices e^A et e^B entraîne celle des matrices A et B.]

Peut-on déduire la formule (8) de la formule (11)?

Remarque 1. Bien que l'on s'intéresse essentiellement — tant que le contraire n'est pas expressément mentionné — aux *matrices réelles* (i.e. dont les éléments sont réels), on sera parfois contraint de faire intervenir les matrices complexes (cf., par exemple, démonstration de la formule (6) ci-dessus). Il importe donc de remarquer que ce qui a été dit plus haut sur les matrices e^A et $\ln A$ vaut pour les matrices aussi bien réelles que complexes.

* * *

On rappelle (cf. leçon I. 14) qu'une matrice réelle A est *orthogonale* si sa transposée $A^\top$ est son inverse, i.e. si $A^\top A = E$.

Les matrices orthogonales d'ordre n forment un groupe que l'on note $O(n)$.

D'une façon générale, si J est une matrice régulière d'ordre n, la matrice A d'ordre n telle que

$$A^\top J A = J \tag{14}$$

s'appelle *matrice J-orthogonale*. (Donc, les matrices E-orthogonales ne sont autres que les matrices orthogonales.) Les matrices J-orthogonales forment un groupe qui est désigné par $O_J(n)$.

De la formule (14) il résulte que le déterminant d'une matrice J-orthogonale A est égal à ± 1. Une matrice J-orthogonale A telle que $\det A = 1$ s'appelle *matrice propre* ou *unimodulaire*.

Les matrices J-orthogonales unimodulaires forment un sous-groupe $SO_J(n)$ du groupe $O_J(n)$ (pour $J = E$, le sous-groupe $SO(n)$).

Si pour une matrice J-orthogonale A on a l'inégalité $|A - E| < 1$ et donc existe la matrice $B = \ln A$, alors

$$e^{JBJ^{-1}} = Je^BJ^{-1} = JAJ^{-1} = (A^\top)^{-1} = (A^{-1})^\top = (e^{-B})^\top = e^{-B^\top}.$$

Donc, si $|B| < \ln 2$ (et, par suite, $|JBJ^{-1}| < \ln 2$ et $|-B^\top| = |B| < \ln 2$), alors $JBJ^{-1} = -B^\top$, i.e.

$$B^\top J = -JB. \tag{15}$$

Réciproquement, si (15) est réalisée, alors

$$(e^B)^\top = e^{B^\top} = e^{-JBJ^{-1}} = Je^{-B}J^{-1} = J(e^B)^{-1}J^{-1},$$

donc, la matrice $A = e^B$ est une matrice J-orthogonale.

Les matrices satisfaisant à la condition (15) sont dites *J-antisymétriques*. (Pour $J = E$ ce sont d'ordinaires matrices antisymétriques.)

La relation (15) étant linéaire en B, *les matrices J-antisymétriques (d'ordre n donné) forment un sous-espace vectoriel de l'espace* $\mathrm{Mat}_n(\mathbb{R})$. Ce sous-espace est désigné par $\mathfrak{so}_J(n)$ (et pour $J = E$, par $\mathfrak{so}(n)$).

* * *

Plaçons-nous dans le cas général.

Soit $\mathcal{G}$ le groupe des matrices réelles d'ordre n (un sous-groupe du groupe $\mathrm{GL}(n) = \mathrm{GL}(n; \mathbb{R})$).

Définition 1. Un groupe $\mathcal{G}$ s'appelle *groupe de Lie de matrices* si l'espace vectoriel $\mathrm{Mat}_n(\mathbb{R})$ contient un sous-espace vectoriel $\mathfrak{g}$ tel que

1° Pour toute matrice $B \in \mathfrak{g}$ la matrice e^B appartient au groupe $\mathcal{G}$.

2° Si $e^B \in \mathcal{G}$ et $|B| < \ln 2$, alors $B \in \mathfrak{g}$.

L'espace vectoriel $\mathfrak{g}$ s'appelle *algèbre de Lie* du groupe $\mathcal{G}$. (Cette dénomination s'explique par le fait que $\mathfrak{g}$ est une algèbre au sens algébrique général par rapport à l'opération $[B_1, B_2] = B_1B_2 - B_2B_1$; cf. plus bas remarque 1 de la leçon 16.)

Donc, d'après les calculs effectués plus haut, *pour toute matrice régulière J le groupe $\mathrm{O}_J(n)$ est un groupe de Lie de matrices d'algèbre de Lie $\mathfrak{so}_J(n)$.*

Le groupe $\mathrm{GL}(n, \mathbb{R})$ est de toute évidence un groupe de Lie de matrices (pour ce groupe, $\mathfrak{g} = \mathrm{Mat}_n(\mathbb{R})$; remarquons à propos que pour cette raison l'espace $\mathrm{Mat}_n(\mathbb{R})$ se désigne aussi par $\mathfrak{gl}(n)$). Est aussi groupe de Lie de matrices le sous-groupe $\mathrm{GL}^+(n)$ du groupe $\mathrm{GL}(n)$, composé des matrices à déterminant strictement positif (pour ce sous-groupe on a aussi $\mathfrak{g} = \mathrm{Mat}_n(\mathbb{R})$).

De la formule (6) il découle immédiatement que dans le cas du groupe $\mathrm{SL}(n) = \mathrm{SL}(n; \mathbb{R})$ des matrices unimodulaires, les propriétés 1° et 2° sont vérifiées par le sous-espace $\mathfrak{sl}(n)$ à $n^2 - 1$ dimensions des matrices B telles que $\mathrm{Tr}\, B = 0$. Donc, *le groupe* $\mathrm{SL}(n)$ *est un groupe de Lie de matrices d'algèbre de Lie* $\mathfrak{sl}(n)$.

L'importance des groupes de Lie de matrices en théorie des variétés différentiables est mise en évidence par la proposition suivante:

Proposition 1. *Tout groupe de Lie de matrices $\mathcal{G}$ est muni de façon naturelle d'une structure de variété différentiable.*

La dimension de cette variété est égale à celle de l'algèbre $\mathfrak{g}$.

Démonstration. Soit U_0 un voisinage de la matrice nulle 0 de $\mathfrak{g}$, composé des matrices $B \in \mathfrak{g}$ telles que $|B| < \ln 2$, et soit $U = \exp U_0$. Alors U est ouvert dans $\mathcal{G}$ (pourquoi?) et l'application $\ln : A \to \ln A$ est une application bijective (voire même un homéomorphisme) de U sur U_0.

En fixant un ordre d'écriture des éléments des matrices de $\mathfrak{g}$ sous forme d'une ligne et en identifiant par là même l'algèbre $\mathfrak{g}$ à l'espace $\mathbb{R}^N$, où $N = \dim \mathfrak{g}$, on peut traiter U_0 comme un sous-ensemble de l'espace $\mathbb{R}^N$. Il est clair que ce sous-ensemble est ouvert dans $\mathbb{R}^N$ et, par conséquent, le couple $(U, \ln)$ est une carte sur $\mathfrak{g}$.

Remarquons que $E \in U$.

Pour toute matrice $C \in \mathcal{G}$, considérons l'ensemble $CU \subset \mathcal{G}$ des matrices de $\mathcal{G}$ de la forme CA, où $A \in U$, et l'application

$$h_C : CU \to U_0$$

de cet ensemble sur l'ensemble $U_0 \subset \mathbb{R}^N$, définie par la formule

$$h_C(CA) = \ln A, \quad A \in U.$$

Il est clair que l'application h_C est bijective, i.e. que le couple (CU, h_C) est une carte sur $\mathcal{G}$. Comme $C = CE \in CU$, les cartes de la forme (CU, h_C) constituent un recouvrement de $\mathcal{G}$. D'autre part, pour deux matrices quelconques $C_1, C_2 \in \mathfrak{g}$, l'application

$$h_{C_2} \circ h_{C_1}^{-1} : h_{C_1}(C_1U \cap C_2U) \to h_{C_2}(C_1U \cap C_2U)$$

est définie par la formule

$$(h_{C_2} \circ h_{C_1}^{-1})(B) = \ln(C_2^{-1}C_1e^B),$$

l'ensemble $h_{C_1}(C_1U \cap C_2U)$ étant composé des matrices $B \in U_2$ telles que $C_2^{-1}C_1e^B \in U$.

Donc, l'ensemble $h_{C_1}(C_1U \cap C_2U)$ est ouvert et l'application $h_{C_2} \circ h_{C_1}^{-1}$ est différentiable. Par conséquent, les cartes (C_1U, h_{C_1}) et (C_2U, h_{C_2}) sont compatibles.

On voit donc que les cartes de la forme (CU, h_C) forment un atlas sur $\mathcal{G}$ et, par conséquent, définissent sur $\mathcal{G}$ une structure de variété différentiable (visiblement de dimension N).

Ceci prouve intégralement la proposition 1. □

Dans la suite on admettra toujours que tout groupe de Lie de matrices est muni de la structure différentiable construite.

Etant un sous-ensemble de l'espace vectoriel $\mathrm{Mat}_n(\mathbb{R})$, le groupe $\mathcal{G}$ hérite sa topologie de $\mathrm{Mat}_n(\mathbb{R})$.

Il est clair que toutes les cartes (CU, h_C) vérifient pour cette topologie les conditions de la proposition 2 de la leçon 7, et donc *la structure différentiable du groupe $\mathcal{G}$ est compatible avec sa topologie.*

On voit donc en particulier que *le groupe* $\mathrm{SL}(n)$ *est une variété différentiable de dimension* $n^2 - 1$.

* * *

De façon analogue, *pour toute matrice régulière J d'ordre n le groupe* $\mathrm{O}_J(n)$ *des matrices J-orthogonales d'ordre n est une variété différentiable.*

La dimension de cette variété (qui est égale à celle de l'algèbre de Lie $\mathfrak{so}_J(n)$) dépend généralement de J.

Par exemple, pour $J = E$, l'algèbre $\mathfrak{so}_E(n) = \mathfrak{so}(n)$ est composée des matrices antisymétriques $B = \| b_j^i \|$ d'ordre n. Puisque dans l'espace vectoriel $\mathrm{Mat}_n(\mathbb{R})$ à n^2 dimensions des matrices d'or-

dre n, les matrices antisymétriques sont mises en évidence par $n(n+1)/2$ conditions indépendantes

$$b_j^i = -b_i^j, \quad \text{où} \quad i, j = 1, \ldots, n \quad \text{et} \quad i \leqslant j,$$

on a

$$\dim \mathfrak{so}(n) = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}.$$

Donc, *le groupe de Lie* $O(n)$ *est de dimension* $\frac{n(n-1)}{2}$.

Un autre exemple important correspond au cas où n est pair ($n = 2m$) et la matrice J est de la forme

$$J = \begin{Vmatrix} 0 & E \\ -E & 0 \end{Vmatrix}, \tag{16}$$

où 0 et E sont les matrices nulle et unité d'ordre m. Dans ce cas la condition de J-antisymétrie de la matrice

$$A = \begin{Vmatrix} A_1 & A_2 \\ A_3 & A_4 \end{Vmatrix} \tag{17}$$

d'ordre n (où A_1, A_2, A_3 et A_4 sont des matrices d'ordre m) est équivalente aux trois conditions suivantes

$$A_1^T = -A_4, \quad A_2^T = A_2, \quad A_3^T = A_3, \tag{18}$$

qui expriment que les matrices A_2 et A_3 sont symétriques et que la matrice A_4 s'exprime en fonction de la matrice A_1 (qui n'est astreinte à aucune condition). Donc, la dimension de l'espace des matrices J-antisymétriques (désigné ici par $\mathfrak{sp}(m; \mathbb{R})$) est égale à

$$m^2 + 2 \cdot \frac{m(m+1)}{2} = m(2m+1) = \frac{n(n+1)}{2}.$$

Les matrices J-orthogonales à la matrice (16) s'appellent *matrices symplectiques*, et le groupe $O_J(n)$ qu'elles forment (et qui est généralement désigné par $\mathrm{Sp}(m; \mathbb{R})$) *groupe symplectique*. D'après le calcul effectué, *le groupe* $\mathrm{Sp}(m; \mathbb{R})$ *est une variété différentiable de dimension* $m(2m+1)$.

Les matrices d'ordre pair $n = 2m$, qui sont à la fois orthogonales et symplectiques, forment un *groupe symplectique orthogonal* noté $\mathrm{Sp}(m; \mathbb{R}) \cap O(2m)$. Ce groupe est un groupe de Lie de matrices d'algèbre de Lie $\mathfrak{sp}(m; \mathbb{R}) \cap \mathfrak{so}(2m)$.

Il est clair qu'une matrice (17) est antisymétrique si et seulement si $A_1^T = -A_1$, $A_2^T = -A_3$ et $A_4^T = -A_4$. Combiné aux conditions (18) ceci nous dit qu'*une matrice* (17) *appartient à l'algèbre de Lie* $\mathfrak{sp}(m; \mathbb{R}) \cap \mathfrak{so}(2m)$ *si et seulement si*

$$A_1^T = -A_1, \quad A_2^T = A_2 \quad \textit{et} \quad A_3 = -A_2, \quad A_4 = A_1.$$

Donc,

$$\dim(\mathfrak{sp}(m;\ \mathbb{R}) \cap \mathfrak{so}(2m)) = \frac{m(m-1)}{2} + \frac{m(m+1)}{2} = m^2,$$

et, par suite, *le groupe* $\mathrm{Sp}(m;\ \mathbb{R}) \cap \mathrm{O}(2m)$ *est une variété différentiable de dimension* m^2.

Supposons que $n = p + q$ et soit

$$J = \begin{Vmatrix} E_p & 0 \\ 0 & -E_q \end{Vmatrix},$$

où E_p et E_q sont les matrices unités d'ordre p et q respectivement.

Les matrices J-orthogonales à la matrice J s'appellent *matrices pseudo-orthogonales de type* (p, q), et le groupe $\mathrm{O}_J(n)$ qu'elles forment est désigné par $\mathrm{O}(p, q)$. (Donc, $\mathrm{O}(n, 0) = \mathrm{O}(0, n) = \mathrm{O}(n)$.) Il est clair que *le groupe* $\mathrm{O}(p, q)$ *est une variété différentiable de dimension* $\frac{n(n-1)}{2}$ (ne dépendant pas de p et q).

* * *

Les analogues complexes des matrices J-orthogonales sont les *matrices J-unitaires* à éléments complexes satisfaisant à la relation

$$\bar{A}^{\mathsf{T}} J A = J \tag{19}$$

(pour $J = E$, ce sont d'ordinaires matrices unitaires; cf. leçon II. 21) et ceux des matrices J-antisymétriques, les *matrices J-antihermitiennes* vérifiant la relation

$$\bar{B}^{\mathsf{T}} J = -JB$$

(pour $J = E$, ce sont d'ordinaires matrices antihermitiennes; cf. leçon II.19).

Les matrices J-unitaires forment un groupe $\mathrm{U}_J(n)$ (pour $J = E$, le groupe $\mathrm{U}(n)$), et les matrices J-antihermitiennes, l'espace vectoriel $\mathfrak{u}_J(n)$ (pour $J = E$, l'espace vectoriel $\mathfrak{u}(n)$).

[Signalons que, bien que l'espace vectoriel $\mathfrak{u}_J(n)$ soit composé de matrices complexes, il n'est stable que pour la multiplication par des nombres réels (la multiplication par i donne une matrice hermitienne d'une matrice antihermitienne), i.e. $\mathfrak{u}_J(n)$ *est un espace vectoriel sur le corps* $\mathbb{R}$. Pour $J = E$, on trouve par un calcul simple que sa dimension est égale à n^2. (La condition d'antihermiticité des matrices n'impose aucune contrainte aux $\frac{n(n-1)}{2}$ nombres complexes situés au-dessus de la diagonale principale et implique aux n éléments diagonaux d'être imaginaires purs.)]

Exercice 3. Montrer que pour toute matrice J-antihermitienne B la matrice e^B est une matrice J-unitaire et que, si la matrice $A = e^B$,

avec $|B| < \ln 2$, est J-unitaire, la matrice B est nécessairement J-antihermitienne.

A titre de généralisation immédiate de la définition, on se servira du terme « groupe de Lie de matrices » pour les groupes $\mathcal{G}$ composés de matrices complexes, pour lesquels existe l'espace vectoriel $\mathfrak{g} \subset \subset \mathrm{Mat}_n(\mathbb{C})$ (*sur le corps* $\mathbb{R}$) jouissant des propriétés 1° et 2° de la définition 1. Aux termes de cette définition, on peut donc dire que $\mathrm{U}_J(n)$ *est un groupe de Lie de matrices d'algèbre de Lie* $\mathfrak{u}_J(n)$. Puisque la proposition 1 est de toute évidence valable aussi pour de tels groupes de Lie de matrices, on voit que *le groupe* $\mathrm{U}_J(n)$ *est une variété différentiable*. Pour $J = E$, cette variété est de dimension n^2.

Exercice 4. Montrer que la correspondance

$$\begin{Vmatrix} A_1 & A_2 \\ A_3 & A_4 \end{Vmatrix} \mapsto A_1 + iA_2$$

est un difféomorphisme (et en même temps un isomorphisme) du groupe symplectique orthogonal $\mathrm{Sp}(m; \mathbb{R}) \cap \mathrm{O}(2m)$ sur le groupe $\mathrm{U}(m)$.

De la relation (19) il résulte que pour toute matrice J-unitaire A, on a

$$|\det A| = 1.$$

Les matrices unitaires A telles que

$$\det A = 1$$

(les *matrices unitaires unimodulaires*) forment un sous-groupe $\mathrm{SU}(n)$ du groupe $\mathrm{U}(n)$. Ce groupe est un groupe de Lie dont l'algèbre de Lie est composée de toutes les matrices antihermitiennes d'ordre n de trace nulle (cette algèbre est généralement désignée par $\mathfrak{su}(n)$). Puisque de toute évidence $\dim \mathfrak{su}(n) = n^2 - 1$, *le groupe* $\mathrm{SU}(n)$ *est une variété différentiable de dimension* $n^2 - 1$.

Un autre sous-groupe intéressant du groupe $\mathrm{U}(n)$ est (pour $n = 2m$ pair) le *groupe symplectique unitaire* $\mathrm{Sp}(m)$ des matrices unitaires d'ordre $n = 2m$ vérifiant la condition de symplecticité (qui de toute évidence a un sens pour les matrices complexes aussi). Ce groupe est un groupe de Lie dont l'algèbre de Lie $\mathfrak{sp}(m)$ est composée des matrices antihermitiennes (17) satisfaisant aux relations (18). Mais la condition d'antihermiticité de la matrice (17) se ramène aux relations

$$\bar{A}_1 = -A_1, \quad \bar{A}_4^{\mathrm{T}} = -A_4, \quad \bar{A}_2^{\mathrm{T}} = -A_3$$

qui combinées aux relations (18), nous donnent

$$A_4 = \bar{A}_1 \quad A_3 = -\bar{A}_2,$$

où A_1 est une matrice antihermitienne (dépendant de m^2 paramètres réels) et A_2 une matrice symétrique (dépendant, en raison de sa

complexité, de $m(m+1)$ paramètres réels). Ceci montre que $\dim \mathfrak{sp}(m) = m^2 + m(m+1) = m(2m+1)$ et, par suite, que *le groupe* $\mathrm{Sp}(m)$ *est une variété différentiable de dimension* $m(2m+1)$.

Notons que $\dim \mathrm{Sp}(m) = \dim \mathrm{Sp}(m; \mathbb{R})$. Néanmoins, *les variétés* $\mathrm{Sp}(m)$ *et* $\mathrm{Sp}(m; \mathbb{R})$ *ne sont pas difféomorphes* (ne serait-ce que parce que l'une est compacte et l'autre non).

* * *

Comme déjà remarqué, la condition de symplecticité (et de façon plus générale la condition de J-orthogonalité) a un sens pour les matrices complexes. Ainsi apparaît *le groupe* $\mathrm{O}_J(n; \mathbb{C})$ *des matrices* J*-orthogonales complexes d'ordre* n (et en particulier *le groupe* $\mathrm{O}(n; \mathbb{C})$ *des matrices orthogonales complexes d'ordre* n et *le groupe* $\mathrm{Sp}(m; \mathbb{C})$ *des matrices symplectiques complexes d'ordre* $n = 2m$). Ce groupe satisfait aux conditions de la définition 1 à seule différence que son algèbre de Lie $\mathfrak{g}$ est composée de matrices complexes et est un espace vectoriel sur le corps $\mathbb{C}$. De tels groupes seront appelés *groupes de Lie de matrices complexes.*

Puisque tout espace vectoriel sur le corps $\mathbb{C}$ est automatiquement un espace vectoriel sur le corps $\mathbb{R}$ (de dimension double), tout groupe de matrices complexes sera un groupe de Lie de matrices (au sens général indiqué plus haut) et donc une variété différentiable. On voit en particulier que *tous les groupes de la forme* $\mathrm{O}_J(n; \mathbb{C})$ (et en particulier les groupes $\mathrm{O}(n; \mathbb{C})$ et $\mathrm{Sp}(m; \mathbb{C})$) *sont des variétés différentiables.* De plus, $\dim \mathrm{O}(n; \mathbb{C}) = n(n-1)$ et $\dim \mathrm{Sp}(m; \mathbb{C}) = 2m(2m+1)$.

On remarquera donc que

$$\mathrm{Sp}(m) = \mathrm{Sp}(m; \mathbb{C}) \cap \mathrm{U}(2m)$$

(l'intersection $\mathrm{Sp}(m; \mathbb{R}) \cap \mathrm{U}(2m)$ n'est autre que le groupe symplectique orthogonal $\mathrm{Sp}(m; \mathbb{R}) \cap \mathrm{O}(2m)$).

* * *

L'exemple des groupes $\mathrm{O}_J(n; \mathbb{C})$ nous suggère d'envisager sous leur forme générale les variétés munies de cartes (U, h) pour lesquelles l'ensemble $h(U)$ est un ensemble ouvert de l'espace $\mathbb{C}^n$ (de telles cartes sont dites *complexes*).

Deux cartes complexes (U, h) et (V, k) s'appellent $\mathbb{C}$*-compatibles* si soit $U \cap V = \varnothing$, soit les ensembles $h(U \cap V)$ et $k(U \cap V)$ sont ouverts dans $\mathbb{C}^n$ et l'application

$$k \circ h^{-1} : h(U \cap V) \to k(U \cap V)$$

est un *difféomorphisme* $\mathbb{C}$*-analytique*, i.e. elle et sa réciproque $h \circ k^{-1}$ s'expriment par des fonctions $\mathbb{C}$-analytiques (on dit encore

holomorphes) qui au voisinage de tout point $z_0 = (z_0^1, \ldots, z_0^n) \in h(U \cap V) \subset \mathbb{C}^n$ se développent en séries entières de $z^1 - z_0^1, \ldots, z^n - z_0^n$.

Un ensemble $\mathscr{X}$ muni d'un atlas de cartes complexes $\mathbb{C}$-compatibles s'appelle *variété* $\mathbb{C}$-*analytique*. Le nombre n s'appelle *dimension complexe* de $\mathscr{X}$ et se note $\dim_{\mathbb{C}} \mathscr{X}$.

L'analogue « complexe » de la proposition 1 affirme que *tout groupe de Lie de matrices complexes est une variété* $\mathbb{C}$-*analytique*. La démonstration reprend *ad litteram* celle de la proposition 1.

[Soulignons que le groupe U (n) n ' e s t p a s un groupe de Lie de matrices complexes; d'une façon générale il est impossible de munir U (n) d'une structure de variété $\mathbb{C}$-analytique.]

Puisque l'espace $\mathbb{C}^n$ s'identifie de façon naturelle avec l'espace $\mathbb{R}^{2n}$ (il suffit de remplacer chaque coordonnée complexe z^k, $k = 1, \ldots, n$, par ses parties réelle et imaginaire) et que les parties réelle et imaginaire de toute fonction $\mathbb{C}$-analytique de $z^1, \ldots, z^n$ sont des fonctions $\mathbb{R}$-analytiques (prouvez-le !) de

$$x^1 = \operatorname{Re} z^1, \ \ldots, \ x^n = \operatorname{Re} z^n, \quad y^1 = \operatorname{Im} z^1, \ \ldots, \ y^n = \operatorname{Im} z^n,$$

toute variété $\mathbb{C}$-analytique $\mathscr{X}$ est *ipso facto* une variété différentiable ($\mathbb{R}$-analytique) de dimension double.

Cette variété s'appelle *décomplexifié* de la variété $\mathscr{X}$ et se note $\mathscr{X}_{\mathbb{R}}$.

* * *

Revenons provisoirement à la topologie générale.

Définition 2. Dans un espace topologique $\mathscr{X}$ on appelle *chemin* une application continue $u : I \to \mathscr{X}$ du segment $I = [0, 1]$ dans $\mathscr{X}$. Le point $u(0)$ s'appelle *origine* et le point $u(1)$, *extrémité* du chemin u. On dit aussi que le chemin u *relie* dans $\mathscr{X}$ le point $p_0 = u(0)$ au point $p_1 = u(1)$.

Les déformations des bases d'un espace vectoriel $\mathscr{V}$ à n dimensions (cf. définition 2, leçon I.12) ne sont autres que des chemins dans la variété de Stiefel $V(n, \mathscr{V}) = \mathrm{GL}(n)$, et les déformations pseudo-orthonormales des bases des espaces pseudo-euclidiens de type (p, q) (définies de façon analogue), autres que des chemins dans le groupe $\mathrm{O}(p, q)$.

Il est aisé de voir (comparer avec la proposition 1 de la leçon I.12) que la relation « être relié par un chemin » est une relation d'équivalence sur $\mathscr{X}$, i.e. est une relation réflexive (tout point $p_0 \in \mathscr{X}$ est relié à lui-même par un *chemin constant* $u : I \to \mathscr{X}$ défini par la formule $u(t) = p_0$, $t \in I$), symétrique (si un chemin $u : I \to \mathscr{X}$ relie p_0 à p_1, le *chemin réciproque* $u^{-1} : I \to \mathscr{X}$ défini par la formule $u^{-1}(t) = u(1 - t)$, $t \in I$, reliera p_1 à p_0) et transitive (si un chemin $u : I \to \mathscr{X}$ relie un point p_0 à un point p_1, et un chemin $v : I \to \mathscr{X}$, le

point p_1 à un point p_2, alors le chemin $w : I \to \mathscr{X}$ défini par la formule

$$w(t) = \begin{cases} u(2t) & \text{si } 0 \leqslant t \leqslant 1/2, \\ v(2t-1) & \text{si } 1/2 \leqslant t \leqslant 1, \end{cases} \tag{20}$$

et appelé, à propos, *produit uv* des chemins u et v reliera p_0 à p_2; montrer que la formule (20) définit une application continue).

Définition 3. Les classes modulo cette relation d'équivalence s'appellent *composantes connexes par arcs* de l'espace topologique $\mathscr{X}$. Un espace $\mathscr{X}$ qui possède une seule composante connexe par arcs, i.e. tel que deux quelconques de ses points sont reliables par un chemin, s'appelle *espace connexe par arcs.*

Les orientations d'un espace vectoriel, traitées comme des classes de bases se déduisant l'une de l'autre par une déformation (cf. théorème 1 de la leçon I.12), et les orientations d'un espace pseudo-euclidien définies de façon analogue ne sont autres que les composantes connexes par arcs des groupes GL (n) et O (p, q) respectivement. On voit donc que *les groupes* GL (n) et O (n) *possèdent chacun deux composantes connexes par arcs, et le groupe* O (p, q), $0 < p < n$, *en a quatre.* La composante du groupe GL (n) qui contient l'unité (on l'appelle plus brièvement *composante de l'unité*) est le groupe $\mathrm{GL}^+(n)$ des matrices à déterminant strictement positif, la composante de l'unité du groupe O (n) est le groupe SO (n), et la composante de l'unité du groupe O (p, q), $0 < p < n$, le groupe noté $\mathrm{O}^{\uparrow}_{+}(p, q)$.

Exercice 5. Montrer que les groupes Sp $(m; \mathbb{R})$, U (n), SU (n) et Sp (m) sont connexes par arcs.

* * *

Il existe une autre notion de connexité parfois plus commode (dont nous nous sommes déjà servis à la leçon 1 en étudiant la notion de ligne simple).

On rappelle (cf. leçon 8) qu'un sous-ensemble A d'un espace topologique $\mathscr{X}$ est *fermé* si son complémentaire $\mathscr{X} \setminus A$ est ouvert. Un sous-ensemble à la fois ouvert et fermé s'appelle *ouvert-fermé.*

Il est clair que l'ensemble vide $\varnothing$ et l'espace $\mathscr{X}$ sont ouverts-fermés. On les appelle *sous-ensembles ouverts-fermés triviaux.* Un exemple de sous-ensemble ouvert-fermé non trivial est le sous-groupe SO (n) du groupe O (n).

Définition 4. On dit qu'un espace topologique est *connexe* s'il ne contient pas de sous-ensembles ouverts-fermés non triviaux. Un sous-ensemble d'un espace topologique est *connexe* s'il l'est pour la topologie induite.

De façon imagée la connexité d'un espace exprime qu'il est constitué d'un seul morceau.

Exercice 6. Montrer que la longue demi-droite d'Alexandrov (cf. remarque 1 de la leçon 10) est un espace topologique connexe.

Définition 5. On dit qu'un sous-ensemble connexe A d'un espace topologique $\mathscr{X}$ est une *composante connexe* de $\mathscr{X}$ s'il est maximal, i.e. si chaque sous-ensemble connexe contenant A est confondu avec A.

On sait du cours d'analyse que *tout segment* $[a, b]$ *de l'axe* $\mathbb{R}$ *est un espace connexe.*

Prouvons-le. Supposons que le segment $[a, b]$ n'est pas connexe, i.e. supposons qu'il existe un ensemble ouvert-fermé non vide $C \subset [a, b]$ distinct de $[a, b]$. Quitte à passer au complémentaire (qui est aussi ouvert-fermé) on peut, sans perdre en généralité, admettre que $a \in C$. Soit T l'ensemble de tous les points $t \in [a, b]$ tels que $[a, t[\subset C$. Puisque l'inclusion $[a, t[\subset C$ implique que $[a, t] \subset C$ (car C est fermé), on en déduit que $T \subset C$. L'ensemble C étant ouvert dans $[a, b]$ (et $a \in C$), il existe un $\varepsilon > 0$ tel que $[a, a + \varepsilon[\subset C$, i.e. $a + \varepsilon \in T$. Donc l'ensemble T n'est pas vide et sa borne supérieure $t_0 = \sup T$ est $> a$. Supposons que $a < t < t_0$. Par définition de la borne supérieure, il existe un nombre $t_1 \in T$ tel que $t < t_1$. Alors $[a, t[\subset [a, t_1[\subset C$ et, par suite, $t \in T$. Ceci prouve que $[a, t_0[\subset C$, i.e. que $t_0 \in T \subset C$. Mais si $t_0 \in C$ et $t_0 < b$, c'est qu'il existe, puisque C est ouvert dans $[a, b]$, un $\varepsilon > 0$ tel que $]t_0 - \varepsilon, t_0 + \varepsilon[\subset C$. Donc, $[a, t_0 + \varepsilon[\subset C$ et, par suite, $t_0 + \varepsilon \in T$. Puisque ceci contredit le fait que $t_0 = \sup T$, l'inégalité $t_0 < b$ est impossible. Donc, $t_0 = b$ et, contrairement à l'hypothèse, $C = [a, b]$. □

En particulier, *le segment I est connexe.*

On en déduit sans peine que *tout espace $\mathscr{X}$ connexe par arcs est connexe.* En effet, si C est un ensemble ouvert-fermé dans $\mathscr{X}$, pour tout chemin $u : I \to \mathscr{X}$ l'ensemble $u^{-1}C$ des points $t \in [0, 1]$ tels que $u(t) \in C$ est ouvert-fermé dans I. Donc, soit $u^{-1}C = \varnothing$, soit $u^{-1}C = I$, puisque le segment I est connexe. Si $C \neq \varnothing$, on peut exclure le premier cas en choisissant l'origine du chemin u dans C, et si $C \neq \mathscr{X}$, on peut exclure le deuxième cas en prenant l'extrémité du chemin u à l'extérieur de C. Donc, si dans $\mathscr{X}$ il existe un ensemble C ouvert-fermé non vide distinct de $\mathscr{X}$, aucun point de C ne peut être relié par un chemin à un point extérieur à C. Donc, $\mathscr{X}$ n'est pas connexe par arcs. □

Il est clair maintenant que *toute composante connexe par arcs est contenue dans une (et visiblement une seule) composante connexe.*

Il est aisé de voir que *l'adhérence $\bar{A}$ d'un sous-ensemble connexe A est connexe.* En effet, si un sous-ensemble non vide C est ouvert-fermé dans $\bar{A}$, son intersection $C \cap A$ avec A n'est pas vide (pourquoi ?) et est ouverte-fermée dans A. Donc, $C \cap A = A$ et, par conséquent, $A \subset C$. L'ensemble C étant fermé aussi bien dans $\bar{A}$ que dans $\mathscr{X}$, on en déduit que $\bar{A} \subset C$ et, par suite, $\bar{A} = C$. □

Il s'ensuit en particulier que *toute composante connexe de l'espace $\mathscr{X}$ est fermée.*

On démontre de façon analogue que *si l'intersection $\bigcap A_\alpha$ d'une famille $\{A_\alpha\}$ d'ensembles connexes n'est pas vide, leur réunion $\bigcup A_\alpha$ est connexe.* En effet, si C est ouvert-fermé dans $\bigcup A_\alpha$ et $C \cap A_\alpha \neq \varnothing$, alors $C \cap A_\alpha = A_\alpha$, donc $A_\alpha \subset C$. Mais si $A_{\alpha_0} \subset C$ pour un indice α_0 au moins, alors $\bigcap A_\alpha \subset C$ et, par suite, $C \cap A_\alpha \neq \varnothing$ pour tout α. Donc, soit $A_\alpha \subset C$ pour tout α, soit $C \cap A_\alpha = \varnothing$ aussi pour tout α. Dans le premier cas, $C = \bigcup A_\alpha$, dans le deuxième, $C = \varnothing$. □

Il en résulte que *pour tout point $p \in \mathscr{X}$ la composante connexe de $\mathscr{X}$ qui contient p est la réunion de tous les sous-ensembles connexes de $\mathscr{X}$ contenant p* et que *les composantes connexes de $\mathscr{X}$ sont deux à deux disjointes.* Donc, *si un espace $\mathscr{X}$ possède un nombre fini de composantes connexes, chacune d'elles est ouverte dans $\mathscr{X}$* (donc est un ensemble ouvert-fermé).

En résumé donc, *tout espace topologique $\mathscr{X}$ est la réunion de composantes connexes deux à deux disjointes dont chacune est un sous-ensemble connexe fermé (et ouvert si ces composantes sont en nombre fini) de l'espace $\mathscr{X}$.*

L'espace $\mathscr{X}$ est connexe si et seulement s'il possède une seule composante connexe.

On dit qu'un espace $\mathscr{X}$ est *totalement discontinu* si chacune de ses composantes contient un seul point. Il est clair qu'*un espace séparé $\mathscr{X}$ tel que* $\operatorname{ind} \mathscr{X} = 0$ (cf. leçon 10) *est totalement discontinu.*

D'autre part, il existe des exemples (assez complexes) d'espaces séparés totalement discontinus (voire même métriques et satisfaisant au deuxième axiome de dénombrabilité) de dimension positive quelconque. Ces espaces ne sont visiblement pas compacts puisqu'on peut prouver (faites-le à titre d'exercice) qu'*un espace compact séparé totalement discontinu $\mathscr{X}$ est de dimension* 0 ($\operatorname{ind} \mathscr{X} = \dim \mathscr{X} = 0$).

* * *

Contrairement aux composantes connexes, les composantes connexes par arcs ne sont généralement ni ouvertes ni fermées. Mais *si l'espace $\mathscr{X}$ est localement connexe par arcs* (chacun de ses points possède un système fondamental de voisinages qui sont tous connexes par arcs), *toute composante connexe par arcs C de l'espace $\mathscr{X}$ est ouverte-fermée.* En effet, si U est un voisinage connexe par arcs d'un point $p \in \overline{C}$, alors la réunion $C \cup U$ est connexe par arcs, puisque $U \cap C$ n'est pas vide, donc est confondue avec C, puisque C est maximale. Donc, $U \subset C$ est, par conséquent, $p \in \operatorname{Int} C$. Ceci prouve que $\overline{C} \subset \operatorname{Int} C$ et, par suite, $\overline{C} = \operatorname{Int} C = C$. Donc, la composante C est ouverte-fermée. □

Il s'ensuit immédiatement que *les composantes connexes par arcs d'un espace localement connexe par arcs sont confondues avec ses composantes connexes* (et de plus que *toutes les composantes connexes d'un espace localement connexe par arcs sont ouvertes-fermées indépendamment de leur nombre*).

En particulier, un espace localement connexe par arcs est connexe si et seulement s'il est connexe par arcs.

Puisque tout point d'une variété topologique (et en particulier différentiable) possède un système fondamental de voisinages homéomorphes à une boule ouverte $\mathring{B}^n$ et que la boule $\mathring{B}^n$ est visiblement connexe par arcs, *toute variété topologique est localement connexe par arcs.*

Donc, dans une variété topologique, les composantes connexes sont confondues avec les composantes connexes par arcs et sont des ensembles ouverts-fermés. (En particulier, *toute composante connexe d'une variété est une variété.*)

On voit que les variétés connexes par arcs ne sont autres que les variétés connexes. Pour cette raison, le terme « connexe par arcs » est remplacé par « connexe » pour les variétés.

* * *

Un chemin $u : I \to \mathscr{X}$ d'une variété différentiable $\mathscr{X}$ est dit *différentiable* s'il est la restriction à I d'une application différentiable $u' :]-\varepsilon, 1 + \varepsilon[\to \mathscr{X}$, où $\varepsilon > 0$. (On rappelle que l'intervalle $] - \varepsilon, 1 + \varepsilon[$ est une variété différentiable.) Si l'application u' est partout différentiable à l'exception d'un nombre fini de points, le chemin u est dit *différentiable par morceaux.*

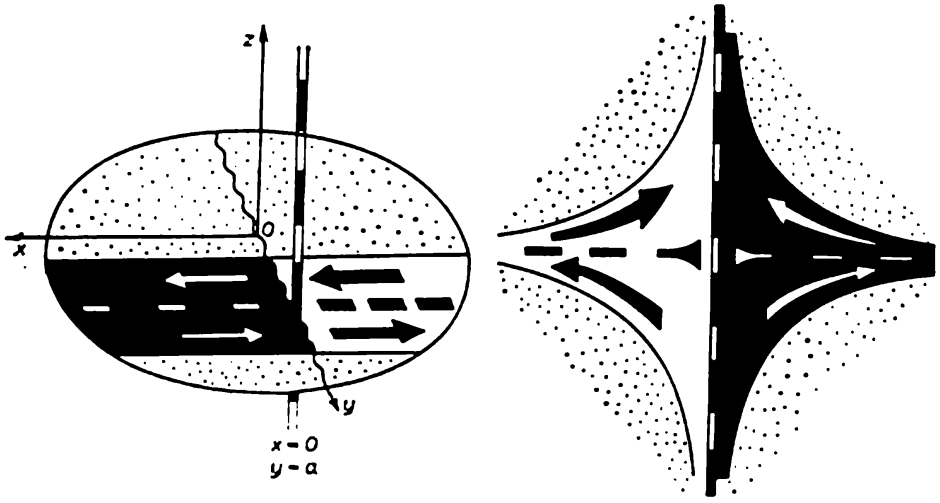
Il est clair que si un ensemble ouvert $U \subset \mathscr{X}$ est difféomorphe à un ensemble ouvert connexe de $\mathbb{R}^n$ (par exemple à une boule), deux quelconques de ses points peuvent être reliés par un chemin différentiable dans U. On en déduit de toute évidence — puisque I est compact — que *deux points quelconques d'une variété différentiable connexe peuvent être reliés par un chemin différentiable par morceaux.*

Exercice 7. Montrer que deux points quelconques d'une variété différentiable connexe peuvent être reliés même par un chemin différentiable.

* * *

Tous les exemples de variétés différentiables connexes de cette leçon satisfont au deuxième axiome de dénombrabilité. Il existe toutefois des variétés différentiables connexes qui ne satisfont pas à cet axiome. Tel est le cas de la longue demi-droite d'Alexandrov (cf. exercice 7 ci-dessus). Citons un exemple plus élémentaire n'utilisant pas les ordinaux (cet exemple appartient à Kalabi et Rosenlicht).

Soit $\mathcal{X}$ le sous-ensemble de l'espace $\mathbb{R}^3$, composé des points (x, y, z) pour lesquels ou bien $x = 0$, ou bien $z = 0$ (la réunion des plans de coordonnées Oyz et Oxy), et soit U_a, $a \in \mathbb{R}$, le sous-ensemble de $\mathcal{X}$ composé des points (x, y, z) pour lesquels ou bien $x \neq 0$, ou bien $y = a$ (la réunion du plan Oxy privé de l'axe des ordonnées $x = 0$, et de la droite $x = 0$, $y = a$). Soit par ailleurs $h_a : U_a \to \mathbb{R}^2$ une



application de U_a sur le plan $\mathbb{R}^2$ muni des coordonnées (u_a, v_a), définie par les formules

$$u_a = x, \quad v_a = \begin{cases} \dfrac{y-a}{x} & \text{si } x \neq 0, \\ z & \text{si } x = 0. \end{cases}$$

Cette application est visiblement bijective (la réciproque envoie le point $(u_a, v_a) \in \mathbb{R}^2$ dans le point $(u_a, a + u_a v_a, v_a) \in U_a$), donc le couple (U_a, h_a) est une carte sur $\mathcal{X}$. Pour deux telles cartes (U_a, h_a) et (U_b, h_b), l'ensemble $U_a \cap U_b$ est le plan Oxy privé de l'axe des ordonnées $x = 0$, les ensembles $h_a(U_a \cap U_b)$ et $h_b(U_a \cap U_b)$ représentent le plan $\mathbb{R}^2$ privé de l'axe des ordonnées $u = 0$, et l'application

$$h_b \circ h_a^{-1} : h_a(U_a \cap U_b) \to h_b(U_a \cap U_b)$$

est définie par les formules

$$u_b = u_a, \quad v_b = v_a + \frac{a-b}{u_a},$$

donc est $\mathbb{R}$-analytique. Ceci exprime que les cartes (U_a, h_a), $a \in \mathbb{R}$, sont $\mathbb{R}$-analytiquement compatibles et, par suite, définissent sur $\mathcal{X}$ une structure de variété différentiable (de classe C^ω) à deux dimensions.

Exercice 8. Montrer que la variété différentiable construite est séparée, connexe et ne satisfait pas au deuxième axiome de dénombrabilité.

Il est évident que cet exemple peut facilement être généralisé à une plus grande dimension. De plus, si l'on remplace $\mathbb{R}$ par $\mathbb{C}$, cette construction nous donne une variété connexe $\mathbb{C}$-analytique à deux dimensions ne satisfaisant pas au deuxième axiome de dénombrabilité.

Il est intéressant de noter que pour $n = 1$ cet exemple est impossible, car, comme l'a prouvé le mathématicien hongrois Rado, *chaque courbe complexe* (une variété connexe $\mathbb{C}$-analytique séparée à une dimension) *satisfait au deuxième axiome de dénombrabilité*.

Remarque 2. La comparaison des définitions montre immédiatement que les courbes complexes ne sont autres que les surfaces riemanniennes abstraites de la théorie des fonctions d'une variable complexe.

LEÇON 12

Vecteurs tangents à une variété différentiable. — Dérivées de fonctions holomorphes. — Vecteurs tangents à des variétés $\mathbb{C}$-analytiques. — Différentielle d'une application différentiable. — Règle en chaîne. — Gradient d'une fonction différentiable. — Théorème des applications étales. — Théorème de changement des coordonnées locales. — Applications localement planes.

Introduisons la notion de vecteur tangent à une variété différentiable $\mathcal{X}$.

Pour comprendre comment on peut le faire, considérons ce cas particulier de la variété $\mathbb{R}^n$ où la notion de vecteur est connue.

Dans des coordonnées arbitraires $x^1, \ldots, x^n$ (éventuellement curvilignes) un vecteur **a** de l'espace affine $\mathbb{R}^n$ est défini en chaque point p_0 de $\mathbb{R}^n$ par n nombres $(a^1, \ldots, a^n)$ (un vecteur de l'espace vectoriel $\mathbb{R}^n$). Comment se transforment ces nombres lorsqu'on passe à d'autres coordonnées $x^{1'}, \ldots, x^{n'}$? (Du cours d'algèbre linéaire on connaît la réponse uniquement dans le cas où les nouvelles coordonnées s'expriment linéairement en fonction des anciennes; mais ici on s'intéresse à la transformation la plus générale des coordonnées, définie par des fonctions différentiables arbitraires $x^{i'} = x^{i'}(x^1, \ldots, x^n)$, $i' = 1, \ldots, n$.) Pour répondre à cette question considérons dans $\mathbb{R}^n$ une courbe différentiable quelconque

$$x^i = x^i(t), \quad i = 1, \ldots, n,$$

passant pour $t = t_0$ par le point p_0 et présentant en ce point un vecteur tangent **a**, i.e. une courbe telle que $a^i = \dot{x}^i(t_0)$ pour tout $i = 1, \ldots, n$ (on désigne la dérivation par rapport à t par un point). Dans les coordonnées $x^{1'}, \ldots, x^{n'}$, cette courbe sera définie par des fonctions $x^{i'}(t) = x^{i'}(x^1(t), \ldots, x^n(t))$, $i' = 1', \ldots, n'$, et le vecteur tangent (i.e. le m ê m e vecteur **a**) aura pour composantes $a^{i'} = \dot{x}^{i'}(t_0)$. Mais d'après la règle de dérivation d'une fonction composée,

$$\dot{x}^{i'}(t) = \frac{\partial x^{i'}(x^1(t), \ldots, x^n(t))}{\partial x^i} \dot{x}^i(t)$$

(comme toujours la sommation est étendue à i), donc

$$a^{i'} = \left(\frac{\partial x^{i'}}{\partial x^i}\right)_0 a^i, \tag{1}$$

où $\left(\frac{\partial x^{i'}}{\partial x^i}\right)_0 = \left(\frac{\partial x^{i'}}{\partial x^i}\right)_{p_0}$ sont les valeurs des dérivées partielles $\frac{\partial x^{i'}}{\partial x^i}$ au point p_0 de coordonnées $x^1(t_0), \ldots, x^n(t_0)$.

Soit maintenant $\mathscr{X}$ une variété différentiable (de classe C^r, $r \geqslant 1$) à n dimensions. Pour tout point p_0 de $\mathscr{X}$ on désignera par $\mathsf{A}(p_0)$ l'ensemble de toutes les cartes (U, h) de $\mathscr{X}$ telles que $p_0 \in U$ (on dit que ces cartes sont *centrées en* p_0). Le calcul que nous venons d'effectuer motive la définition suivante:

Définition 1. On appelle *vecteur tangent à la variété* $\mathscr{X}$ en un point p_0 une application

$$A : \mathsf{A}(p_0) \to \mathbb{R}^n \tag{2}$$

telle que pour des cartes arbitraires $(U, h) = (U, x^1, \ldots, x^n)$ et $(U', h') = (U', x^{1'}, \ldots, x^{n'})$ de $\mathsf{A}(p_0)$ les vecteurs $A(U, h) = (a^1, \ldots, a^n)$ et $A(U', h') = (a^{1'}, \ldots, a^{n'})$ de l'espace vectoriel $\mathbb{R}^n$ sont reliés par la formule (1), i.e. par la formule

$$A(U', h') = \left(\frac{\partial h'}{\partial h}\right)_0 A(U, h), \tag{3}$$

où $\left(\frac{\partial h'}{\partial h}\right)_0$ est un opérateur linéaire dans $\mathbb{R}^n$ de matrice $\left\|\left(\frac{\partial x^{i'}}{\partial x^i}\right)_0\right\|$.

Les composantes $a^1, \ldots, a^n$ du vecteur $A(U, h) \in \mathbb{R}^n$ s'appellent *composantes du vecteur A dans la carte* (U, h) (ou *dans les coordonnées locales* $x^1, \ldots, x^n$). Pour alléger les formules, on met généralement l'égalité $A(U, h) = (a^1, \ldots, a^n)$ sous la forme suivante:

$$A = (a^1, \ldots, a^n) \text{ dans } (U, h).$$

Lorsque la carte (U, h) est fixée, la mention « dans (U, h) » est en principe omise.

L'ensemble de tous les vecteurs tangents à la variété $\mathscr{X}$ en p_0 se désigne par $\mathrm{T}_{p_0}\mathscr{X}$ et s'appelle *espace tangent* à la variété $\mathscr{X}$ en p_0. Cet ensemble est un espace vectoriel sur le corps $\mathbb{R}$ par rapport aux opérations linéaires définies par les formules

$$(A + B)(U, h) = A(U, h) + B(U, h), \quad (\lambda A)(U, h) = \lambda A(U, h),$$

où $A, B \in \mathrm{T}_{p_0}\mathscr{X}$, $\lambda \in \mathbb{R}$ et $(U, h) \in \mathsf{A}(p)$. [Les appartenances $A + B \in \mathrm{T}_{p_0}\mathscr{X}$ et $\lambda A \in \mathrm{T}_{p_0}\mathscr{X}$ sont assurées par la linéarité des opérateurs $\left(\frac{\partial h'}{\partial h}\right)_0$.]

Donc, par définition, si $A = (a^1, \ldots, a^n)$ et $B = (b^1, \ldots, b^n)$ dans (U, h), alors $A + B = (a^1 + b^1, \ldots, a^n + b^n)$ et $\lambda A = (\lambda a^1, \ldots, \lambda a^n)$ dans (U, h). Donc, pour toute carte (U, h) la correspondance

$$A \mapsto A(U, h)$$

définit une application linéaire

$$\mathbf{T}_{p_0}\mathscr{X} \to \mathbb{R}^n. \tag{4}$$

Il est aisé de voir que *l'application* (4) *est un isomorphisme*. En effet, si $A(U, h) = 0$, la formule (3) nous dit que $A(U', h') = 0$ pour toute carte $(U', h') \in \mathsf{A}(p_0)$, i.e. $A = 0$. Ceci exprime que l'application (4) est un monomorphisme. D'autre part, en posant pour tout vecteur $\mathbf{a} = (a^1, \ldots, a^n) \in \mathbb{R}^n$ et toute carte $(U', h') \in \mathsf{A}(p_0)$

$$A(U', h') = \left(\frac{\partial h'}{\partial h}\right)_0 \mathbf{a},$$

on obtient une application $A : \mathsf{A}(p_0) \to \mathbb{R}^n$ telle que $A(U, h) = \mathbf{a}$. Donc, pour prouver que l'application (4) est un épimorphisme (donc un isomorphisme), il suffit de montrer que $A \in \mathbf{T}_{p_0}\mathscr{X}$, i.e. que

$$A(U'', h'') = \left(\frac{\partial h''}{\partial h'}\right)_0 A(U', h')$$

pour deux cartes quelconques (U', h'), $(U'', h'') \in \mathsf{A}(p_0)$. Mais par définition

$$A(U'', h'') = \left(\frac{\partial h''}{\partial h}\right)_0 \mathbf{a} \quad \text{et} \quad A(U', h') = \left(\frac{\partial h'}{\partial h}\right)_0 \mathbf{a},$$

et d'après la règle de dérivation d'une fonction composée

$$\left(\frac{\partial x^{i''}}{\partial x^i}\right)_0 = \left(\frac{\partial x^{i''}}{\partial x^{i'}}\right)_0 \left(\frac{\partial x^{i'}}{\partial x^i}\right)_0,$$

i.e. $\left(\frac{\partial h''}{\partial h}\right)_0 = \left(\frac{\partial h''}{\partial h'}\right)_0 \left(\frac{\partial h'}{\partial h}\right)_0$. Donc,

$$A(U'', h'') = \left(\frac{\partial h''}{\partial h}\right)_0 \mathbf{a} = \left(\frac{\partial h''}{\partial h'}\right)_0 A(U', h'). \quad \square$$

Donc, $\mathbf{T}_{p_0}\mathscr{X}$ *est un espace vectoriel de dimension* n:

$$\dim \mathbf{T}_{p_0}\mathscr{X} = \dim \mathscr{X}.$$

Les vecteurs qui se transforment par l'isomorphisme (4) en la base canonique $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de l'espace $\mathbb{R}^n$ sont désignés par

$$\left(\frac{\partial}{\partial x^1}\right)_{p_0}, \ldots, \left(\frac{\partial}{\partial x^n}\right)_{p_0}. \tag{5}$$

Ils forment une base de l'espace $\mathbf{T}_{p_0}\mathscr{X}$, les coordonnées d'un vecteur $A \in \mathbf{T}_{p_0}\mathscr{X}$ dans cette base étant ses coordonnées dans la carte (U, h): si $A = (a^1, \ldots, a^n)$ dans (U, h), alors

$$A = a^i \left(\frac{\partial}{\partial x^i}\right)_{p_0} \tag{6}$$

et réciproquement.

Lorsque $\mathscr{X}$ est l'espace $\mathbb{R}^n$, il existe parmi les isomorphismes (4) un qui est privilégié: c'est celui qui est associé à la carte $(\mathbb{R}^n, \mathrm{id})$ et qui permet d'identifier l'espace $\mathbf{T}_{p_0}\mathbb{R}^n$ à l'espace $\mathbb{R}^n$. Donc,

$$\mathbf{T}_{p_0}\mathbb{R}^n = \mathbb{R}^n \tag{7}$$

pour tout point $p_0 \in \mathbb{R}^n$, ce qui est parfaitement en accord avec le début de l'exposé.

Puisque pour tout point p_0 d'une sous-variété ouverte U d'une variété $\mathscr{X}$ l'espace $\mathbf{T}_{p_0}U$ s'identifie de façon naturelle à $\mathbf{T}_{p_0}\mathscr{X}$:

$$\mathbf{T}_{p_0}U = \mathbf{T}_{p_0}\mathscr{X},$$

on obtient en particulier

$$\mathbf{T}_{p_0}U = \mathbb{R}^n \tag{8}$$

pour tout point $p_0 \in U$ d'un ensemble ouvert quelconque U de $\mathbb{R}^n$.

On se servira en permanence de ces identifications sans le spécifier expressément à chaque fois.

De façon plus générale, si $\mathscr{X}$ est un espace vectoriel $\mathscr{V}$ (ou un ensemble ouvert dans $\mathscr{V}$), parmi les isomorphismes (4) il en existe qui sont associés aux cartes de la forme $(\mathscr{V}, h)$, où h est un isomorphisme de coordonnées $\mathscr{V} \to \mathbb{R}^n$ associé à une base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ dans $\mathscr{V}$. Puisque pour deux tels isomorphismes h et h' l'opérateur linéaire $\left(\frac{\partial h'}{\partial h}\right)_0$, $p_0 \in \mathscr{V}$, est visiblement confondu avec l'opérateur $h' \circ h^{-1}$ (et en particulier ne dépend pas de p_0), le composé $\mathbf{T}_{p_0}\mathscr{X} \to \mathscr{V}$ de l'isomorphisme (4) associé à la carte $(\mathscr{V}, h)$ et de l'isomorphisme $h^{-1}: \mathbb{R}^n \to \mathscr{V}$ est le même pour tous les h. Donc, $\mathbf{T}_{p_0}\mathscr{V}$ *s'identifie de façon naturelle à* $\mathscr{V}$:

$$\mathbf{T}_{p_0}\mathscr{V} = \mathscr{V} \text{ pour tout } p_0 \in \mathscr{V}.$$

On remarquera que, dans cette identification, à la base (5) est associée la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

Si $\mathscr{X}$ est un espace affine $\mathscr{A}$, pour tout point $p_0 \in \mathscr{A}$ l'espace tangent $\mathbf{T}_{p_0}\mathscr{A}$ s'identifie de façon naturelle à l'espace vectoriel $\mathscr{V}$ auquel $\mathscr{A}$ est attaché (donc, $\mathbf{T}_{p_0}\mathscr{A}$ est de nouveau le même pour tous les p_0). Du reste, il est commode d'introduire une dépendance forcée par rapport à p_0 et d'identifier $\mathbf{T}_{p_0}\mathscr{A}$ à l'espace $\mathscr{A}$ dans lequel p_0 sera pris pour origine, autrement dit admettre que p_0 est l'origine de tous les vecteurs de $\mathbf{T}_{p_0}\mathscr{A}$.

Lorsque $\mathscr{X}$ est une surface élémentaire dans l'espace affine $\mathscr{A}$, l'espace tangent $\mathbf{T}_{p_0}\mathscr{X}$ s'identifie de façon naturelle à l'espace tangent de la leçon 3 (qui est un sous-espace de l'espace $\mathscr{V}$). Plus exactement, si u, v sont des coordonnées locales sur $\mathscr{X}$ correspondant à une paramétrisation $\mathbf{r} = \mathbf{r}(u, v)$ de la surface $\mathscr{X}$, les vecteurs de base correspondants $\left(\frac{\partial}{\partial u}\right)_{p_0}$, $\left(\frac{\partial}{\partial v}\right)_{p_0}$ de l'espace $\mathbf{T}_{p_0}\mathscr{X}$ s'iden-

tifient aux vecteurs $\mathbf{r}_{u_0}$ et $\mathbf{r}_{v_0}$ de l'espace $\mathcal{T}'$. Puisque lors du changement de coordonnées les vecteurs $\left(\frac{\partial}{\partial u}\right)_{p_0}$, $\left(\frac{\partial}{\partial v}\right)_{p_0}$ se transforment d'après les mêmes formules que les vecteurs $\mathbf{r}_{u_0}$ et $\mathbf{r}_{v_0}$, cette identification de l'espace $\mathbf{T}_{p_0}\mathcal{X}$ au sous-espace de l'espace $\mathcal{T}'$ est indépendante du choix de la paramétrisation $\mathbf{r}=\mathbf{r}(u, v)$.

* * *

La notion d'espace tangent s'étend immédiatement aux variétés $\mathbb{C}$-analytiques (cf. leçon 11). Pour réaliser cette extension de façon élégante, nous aurons besoin de quelques notations d'analyse complexe que nous rappellerons en premier lieu.

Pour une fonction complexe $w(z) = u(x, y) + iv(x, y)$ de la variable complexe $z = x + iy$, on posera

$$\frac{\partial w}{\partial z}=\frac{1}{2}\left(\frac{\partial w}{\partial x}-i\frac{\partial w}{\partial y}\right)=\frac{1}{2}\left[\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)+i\left(\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}\right)\right] \tag{9}$$

et

$$\frac{\partial w}{\partial \bar z}=\frac{1}{2}\left(\frac{\partial w}{\partial x}+i\frac{\partial w}{\partial y}\right)=\frac{1}{2}\left[\left(\frac{\partial u}{\partial x}-\frac{\partial v}{\partial y}\right)+i\left(\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\right)\right].$$

[Ces notations sont justifiées par la formule

$$dw=\frac{\partial w}{\partial z}\,dz+\frac{\partial w}{\partial \bar z}\,d\bar z,$$

où $dz = dx + i\,dy$ et $d\bar z = dx - i\,dy$.]

Les équations de Cauchy-Riemann

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x}=-\frac{\partial u}{\partial y},$$

qui caractérisent les fonctions holomorphes se représentent maintenant par une seule égalité

$$\frac{\partial w}{\partial \bar z}=0.$$

[Les fonctions holomorphes peuvent donc être traitées comme des fonctions « indépendantes de $\bar z$ » ! Pour cette raison, les fonctions de z s'écrivent parfois sous la forme $w = w(z, \bar z)$, et la notation $w = w(z)$ n'est utilisée que pour les fonctions holomorphes.]

On remarquera que, si une fonction $w = w(z)$ est holomorphe, les équations de Cauchy-Riemann nous donnent

$$\frac{\partial w}{\partial z}=\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x},$$

i.e. *la dérivée* $\frac{\partial w}{\partial z}$ *d'une fonction holomorphe* $w=w(z)$ *est confondue avec la dérivée ordinaire* $\frac{\partial w}{\partial x}$. De façon analogue,

$$\frac{\partial w}{\partial z}=\frac{1}{i}\frac{\partial w}{\partial y}.$$

Il s'ensuit que pour toutes fonctions holomorphes $w = w(z)$ et $\zeta = \zeta(w)$, la formule

$$\frac{\partial \zeta}{\partial x} = \frac{\partial \zeta}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial \zeta}{\partial v}\frac{\partial v}{\partial x}$$

de dérivation de la fonction composée $\zeta = \zeta(w(z))$ peut être mise sous la forme suivante :

$$\frac{\partial \zeta}{\partial z} = \frac{\partial \zeta}{\partial w}\frac{\partial w}{\partial z}.$$

Ceci exprime que *les dérivées* (9) *des fonctions holomorphes sont justiciables de la règle de dérivation d'une fonction composée.*

Les dérivées partielles $\frac{\partial w}{\partial z^j}$ et $\frac{\partial w}{\partial \bar{z}^j}$, $1 \leqslant j \leqslant n$, d'une fonction complexe $w(\mathbf{z}) = w(z^1, \ldots, z^n)$ de n variables complexes se définissent de façon analogue, et une fonction w est holomorphe si et seulement si

$$\frac{\partial w}{\partial \bar{z}^1} = 0, \ldots, \frac{\partial w}{\partial \bar{z}^n} = 0.$$

Chaque application $w : U \to \mathbb{C}^m$, où $U \subset \mathbb{C}^n$, est définie par m fonctions complexes

(10) $$w^1 = w^1(\mathbf{z}), \ldots, w^m = w^m(\mathbf{z}), \quad \mathbf{z} = (z^1, \ldots, z^m) \in U,$$

de n variables complexes $z^1, \ldots, z^n$ définies sur U. En identifiant $\mathbb{C}^n$ à $\mathbb{R}^{2n} = \mathbb{R}^n \oplus \mathbb{R}^n$ à l'aide de la correspondance

$$\mathbf{z} = (z^1, \ldots, z^n) \Longleftrightarrow (x^1, \ldots, x^n, y^1, \ldots, y^n) = (\mathbf{x}, \mathbf{y}),$$

et $\mathbb{C}^m$ à $\mathbb{R}^{2m} = \mathbb{R}^m \oplus \mathbb{R}^m$ à l'aide de la correspondance

$$\mathbf{w} = (w^1, \ldots, w^m) \Longleftrightarrow (u^1, \ldots, u^m, v^1, \ldots, v^m) = (\mathbf{u}, \mathbf{v}),$$

où

$$x^1 = \operatorname{Re} z^1, \ldots, x^n = \operatorname{Re} z^n, \quad u^1 = \operatorname{Re} w^1, \ldots, u^m = \operatorname{Re} w^m,$$
$$y^1 = \operatorname{Im} z^1, \ldots, y^n = \operatorname{Im} z^n, \quad v^1 = \operatorname{Im} w^1, \ldots, v^m = \operatorname{Im} w^m,$$

on peut traiter w comme une application de $U \subset \mathbb{R}^{2n}$ dans $\mathbb{R}^{2m}$, définie par $2m$ fonctions réelles

$$u^j = u^j(\mathbf{x}, \mathbf{y}), \quad v^j = v^j(\mathbf{x}, \mathbf{y}), \quad 1 \leqslant j \leqslant n.$$

Si cette application est différentiable, sa *matrice jacobienne* (la matrice des dérivées partielles) est de la forme

$$\begin{Vmatrix} \frac{\partial \mathbf{u}}{\partial \mathbf{x}} & \frac{\partial \mathbf{u}}{\partial \mathbf{y}} \\ \frac{\partial \mathbf{v}}{\partial \mathbf{x}} & \frac{\partial \mathbf{v}}{\partial \mathbf{y}} \end{Vmatrix},$$

où

$$\frac{\partial \mathbf{u}}{\partial \mathbf{x}} = \left\| \frac{\partial u^j}{\partial x^k} \right\|, \quad \frac{\partial \mathbf{u}}{\partial \mathbf{y}} = \left\| \frac{\partial u^j}{\partial y^k} \right\|, \quad \frac{\partial \mathbf{v}}{\partial \mathbf{x}} = \left\| \frac{\partial v^j}{\partial x^k} \right\|, \quad \frac{\partial \mathbf{v}}{\partial \mathbf{y}} = \left\| \frac{\partial v^j}{\partial y^k} \right\|.$$

Donc,

$$\left\| \begin{matrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{matrix} \right\| = \left\| \begin{matrix} \frac{E_n}{2} & \frac{E_n}{2} \\ \frac{E_n}{2} & -\frac{E_n}{2} \end{matrix} \right\| \cdot \left\| \begin{matrix} \frac{\partial w}{\partial z} & \frac{\partial w}{\partial \bar{z}} \\ \frac{\overline{\partial w}}{\partial \bar{z}} & \frac{\overline{\partial w}}{\partial z} \end{matrix} \right\| \cdot \left\| \begin{matrix} E_m & iE_m \\ E_m & -iE_m \end{matrix} \right\|,$$

où E_n et E_m sont les matrices unités d'ordre n et m respectivement et

$$\frac{\partial w}{\partial z} = \left\| \frac{\partial w^j}{\partial z^k} \right\|, \quad \frac{\partial w}{\partial \bar{z}} = \left\| \frac{\partial w^j}{\partial \bar{z}^k} \right\|.$$

Lorsque *l'application w est holomorphe* (i.e. les fonctions (10) sont holomorphes, et donc $\frac{\partial w}{\partial \bar{z}} = 0$), il en résulte que

$$(11) \qquad \left\| \begin{matrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{matrix} \right\| = \left\| \begin{matrix} \frac{E_n}{2} & \frac{E_n}{2} \\ \frac{E_n}{2i} & -\frac{E_n}{2i} \end{matrix} \right\| \cdot \left\| \begin{matrix} \frac{\partial w}{\partial z} & 0 \\ 0 & \frac{\overline{\partial w}}{\partial z} \end{matrix} \right\| \cdot \left\| \begin{matrix} E_m & iE_m \\ E_m & -iE_m \end{matrix} \right\|.$$

Pour $n = m$, il s'ensuit en particulier que *le jacobien D_w de l'application holomorphe w: $U \to \mathbb{C}^n$ s'exprime par la formule*

$$D_w = \left| \frac{\partial w}{\partial z} \right|^2,$$

et donc est > 0.

D'autre part, puisque

$$\left\| \begin{matrix} E & iE \\ E & -iE \end{matrix} \right\| \cdot \left\| \begin{matrix} \frac{E}{2} & \frac{E}{2} \\ \frac{E}{2i} & -\frac{E}{2i} \end{matrix} \right\| = \left\| \begin{matrix} E & 0 \\ 0 & E \end{matrix} \right\|,$$

il s'ensuit aussi que (pour tous n et m) les matrices

$$\left\| \begin{matrix} \frac{\partial w}{\partial z} & \frac{\partial w}{\partial \bar{z}} \\ \frac{\overline{\partial w}}{\partial \bar{z}} & \frac{\overline{\partial w}}{\partial z} \end{matrix} \right\|$$

des applications se multiplient lors de la composition de ces applications (car ceci est valable pour les matrices jacobiennes).

On en déduit que *la composition d'applications holomorphes multiplie leurs matrices* $\frac{\partial w}{\partial z}$: si $w^k = w^k(z)$ et $\zeta^l = \zeta^l(w)$, alors

$$\frac{\partial \zeta^l}{\partial z^j} = \frac{\partial \zeta^l}{\partial w^k} \cdot \frac{\partial w^k}{\partial z^j}.$$

On voit donc que *les dérivées* (9) *de fonctions holomorphes sont justiciables des mêmes règles formelles que les dérivées partielles ordinaires des fonctions réelles différentiables.*

* * *

Supposons maintenant que $\mathscr{X}$ est une variété $\mathbb{C}$-analytique et soit $p_0 \in \mathscr{X}$. Soit par ailleurs $\mathsf{A}(p_0)$ l'ensemble des cartes complexes (U, h) de la variété $\mathscr{X}$ telles que $p_0 \in U$. Par définition, pour toutes cartes $(U, h) = (U, z^1, \ldots, z^n)$, $(U', h') = (U', z^{1'}, \ldots, z^{n'})$ de $\mathsf{A}(p_0)$, l'application $h' \circ h^{-1}$ est définie par des fonctions holomorphes

(12) $\quad z^{1'} = z^{1'}(\mathbf{z}), \ \ldots, \ z^{n'} = z^{n'}(\mathbf{z}), \quad$ où $\quad \mathbf{z} = (z^1, \ldots, z^n)$.

Soit $\left(\frac{\partial h'}{\partial h}\right)_0$ un opérateur linéaire dans $\mathbb{C}^n$ de matrice

$$\left(\frac{\partial \mathbf{z}'}{\partial \mathbf{z}}\right)_0 = \left\| \left(\frac{\partial z^{j'}}{\partial z^j}\right)_0 \right\|, \quad j, j' = 1, \ldots, n,$$

où $\left(\frac{\partial z^{j'}}{\partial z^j}\right)_0$ sont les valeurs prises par les dérivées $\frac{\partial z^{j'}}{\partial z^j}$ des fonctions (12) au point p_0. Par analogie avec la définition 1, on appellera *vecteur tangent* à la variété $\mathscr{X}$ au point p_0 une application

(13) $$C : \mathsf{A}(p_0) \to \mathbb{C}^n$$

telle que pour toutes cartes (U, h) et (U', h') de $\mathsf{A}(p_0)$

(14) $$C(U', h') = \left(\frac{\partial h'}{\partial h}\right)_0 C(U, h).$$

Ces vecteurs forment un espace vectoriel $\mathrm{T}_{p_0}\mathscr{X}$ sur le corps $\mathbb{C}$ de dimension $n = \dim_{\mathbb{C}} \mathscr{X}$, et de plus pour toute carte (U, h) de $\mathsf{A}(p_0)$ l'application

(15) $$\mathrm{T}_{p_0}\mathscr{X} \to \mathbb{C}^n, \quad C \mapsto C(U, h),$$

est un isomorphisme. Les vecteurs se transformant par l'isomorphisme (15) en la base canonique de $\mathbb{C}^n$ sont désignés par

$$\left(\frac{\partial}{\partial z^1}\right)_{p_0}, \ \ldots, \ \left(\frac{\partial}{\partial z^n}\right)_{p_0}.$$

Ils forment une base de l'espace $\mathrm{T}_{p_0}\mathscr{X}$ qui est telle que l'égalité $C(U, h) = (c^1, \ldots, c^n)$ équivaut à l'égalité $C = c^j \left(\frac{\partial}{\partial z^j}\right)_{p_0}$.

Il est intéressant de comparer l'espace complexe à n dimensions $\mathrm{T}_{p_0}\mathscr{X}$ avec l'espace réel $\mathrm{T}_{p_0}\mathscr{X}_{\mathbb{R}}$ à $2n$ dimensions, où $\mathscr{X}_{\mathbb{R}}$ est le décomplexifié de la variété $\mathscr{X}$ (cf. leçon 11).

Pour toute carte complexe (U, h), désignons par $h_{\mathbb{R}}$ l'application h traitée comme une application dans $\mathbb{R}^{2n}$, et par $\mathsf{A}(p_0)_{\mathbb{R}}$ l'ensemble des cartes de la forme $(U, h_{\mathbb{R}})$, où $(U, h) \in \mathsf{A}(p_0)$. Par

définition, l'ensemble $\mathsf{A}(p_0)_{\mathbb{R}}$ est un sous-ensemble de l'ensemble $\mathsf{A}_{\mathbb{R}}(p_0)$ des cartes de la variété $\mathscr{X}_{\mathbb{R}}$ au point p_0, et de plus, bien que $\mathsf{A}(p_0)_{\mathbb{R}} \neq \mathsf{A}_{\mathbb{R}}(p_0)$, la restriction des applications $\mathsf{A}_{\mathbb{R}}(p_0) \to \mathbb{R}^{2n}$ à $\mathsf{A}(p_0)_{\mathbb{R}}$ permet de toute évidence d'identifier chaque vecteur tangent $\mathsf{A}_{\mathbb{R}}(p_0) \to \mathbb{R}^{2n}$ à la variété $\mathscr{X}_{\mathbb{R}}$ en p_0 aux applications

$$A : \mathsf{A}(p_0)_{\mathbb{R}} \to \mathbb{R}^{2n} \tag{16}$$

vérifiant pour toutes cartes (U, h) et (U', h') de $\mathsf{A}(p_0)$ la relation

$$A(U', h'_{\mathbb{R}}) = \left(\frac{\partial h'_{\mathbb{R}}}{\partial h_{\mathbb{R}}}\right)_0 A(U, h_{\mathbb{R}}). \tag{17}$$

D'autre part, chaque vecteur tangent (13) à la variété $\mathscr{X}$ en p_0 sera traité en vertu de l'identification $\mathbb{C}^n = \mathbb{R}^{2n}$ comme une application $C_{\mathbb{R}} : \mathsf{A}(p_0)_{\mathbb{R}} \to \mathbb{R}^{2n}$. Ceci étant, pour toutes cartes (U, h) et (U', h') de $\mathsf{A}(p_0)$, on aura

$$C_{\mathbb{R}}(U', h'_{\mathbb{R}}) = \left(\frac{\partial h'}{\partial h}\right)_0^{\mathbb{R}} C_{\mathbb{R}}(U, h), \tag{18}$$

où $\left(\frac{\partial h'}{\partial h}\right)_0^{\mathbb{R}}$ est l'opérateur $\left(\frac{\partial h'}{\partial h}\right)_0$ traité comme un opérateur dans $\mathbb{R}^{2n}$.

Mais il est clair que, si un opérateur linéaire dans $\mathbb{C}^n$ est défini (dans la base canonique) par une matrice $C = A + iB$, ce même opérateur traité comme un opérateur dans $\mathbb{R}^{2n}$ sera défini par la matrice

$$\begin{Vmatrix} A & -B \\ B & A \end{Vmatrix}.$$

Puisque la matrice de l'opérateur $\left(\frac{\partial h'}{\partial h}\right)_0$ est la matrice

$$\left(\frac{\partial z'}{\partial z}\right)_0 = \left(\frac{\partial x'}{\partial x}\right)_0 + i\left(\frac{\partial y'}{\partial x}\right)_0,$$

il s'ensuit que l'opérateur $\left(\frac{\partial h'}{\partial h}\right)_0^{\mathbb{R}}$ est défini par la matrice

$$\begin{Vmatrix} \left(\frac{\partial x'}{\partial x}\right)_0 & -\left(\frac{\partial y'}{\partial x}\right)_0 \\ \left(\frac{\partial y'}{\partial x}\right)_0 & \left(\frac{\partial x'}{\partial x}\right)_0 \end{Vmatrix},$$

qui, en vertu des équations de Cauchy-Riemann, est confondue avec la matrice

$$\begin{Vmatrix} \left(\frac{\partial x'}{\partial x}\right)_0 & \left(\frac{\partial x'}{\partial y}\right)_0 \\ \left(\frac{\partial y'}{\partial x}\right)_0 & \left(\frac{\partial y'}{\partial y}\right)_0 \end{Vmatrix}$$

de l'opérateur $\left(\frac{\partial h'_{\mathbb{R}}}{\partial h_{\mathbb{R}}}\right)_0$. Ceci montre que

$$\left(\frac{\partial h'}{\partial h}\right)_0^{\mathbb{R}} = \left(\frac{\partial h'_{\mathbb{R}}}{\partial h_{\mathbb{R}}}\right)_0,$$

donc, la formule (18) est confondue avec la formule (17) (pour $A = C_{\mathbb{R}}$). Par conséquent, $C_{\mathbb{R}}$ est un vecteur tangent à la variété $\mathscr{X}_{\mathbb{R}}$ en p_0.

L'application

$$\mathrm{T}_{p_0}\mathscr{X} \to \mathrm{T}_{p_0}\mathscr{X}_{\mathbb{R}}, \quad C \mapsto C_{\mathbb{R}},$$

est visiblement bijective et préserve les sommes et les produits par des réels. En d'autres termes, c'est un isomorphisme de l'espace $\mathrm{T}_{p_0}\mathscr{X}$, traité en raison de l'inclusion $\mathbb{R} \subset \mathbb{C}$ comme un espace vectoriel sur le corps $\mathbb{R}$, sur l'espace $\mathrm{T}_{p_0}\mathscr{X}_{\mathbb{R}}$. En désignant cet espace sur le corps $\mathbb{R}$ par $(\mathrm{T}_{p_0}\mathscr{X})_{\mathbb{R}}$, *on obtient* par conséquent *l'identification naturelle*

$$(\mathrm{T}_{p_0}\mathscr{X})_{\mathbb{R}} = \mathrm{T}_{p_0}\mathscr{X}_{\mathbb{R}}. \tag{19}$$

En ce sens les variétés $\mathscr{X}$ et $\mathscr{X}_{\mathbb{R}}$ possèdent les mêmes vecteurs tangents.

La base

$$\left(\frac{\partial}{\partial z^1}\right)_{p_0}, \ldots, \left(\frac{\partial}{\partial z^n}\right)_{p_0}$$

de l'espace $\mathrm{T}_{p_0}\mathscr{X}$ associée à la carte (U, h) de $\mathsf{A}(p_0)$ engendre visiblement la base

$$\left(\frac{\partial}{\partial z^1}\right)_{p_0}, \ldots, \left(\frac{\partial}{\partial z^n}\right)_{p_0}, \quad J\left(\frac{\partial}{\partial z^1}\right)_{p_0}, \ldots, J\left(\frac{\partial}{\partial z^n}\right)_{p_0} \tag{20}$$

de l'espace $(\mathrm{T}_p\mathscr{X}_0)_{\mathbb{R}}$, où J est un opérateur de structure complexe (la multiplication par i). Il est aisé de voir qu'*en vertu de l'identification* (19), *la base* (20) *est confondue avec la base*

$$\left(\frac{\partial}{\partial x^1}\right)_{p_0}, \ldots, \left(\frac{\partial}{\partial x^n}\right)_{p_0}, \quad \left(\frac{\partial}{\partial y^1}\right)_{p_0}, \ldots, \left(\frac{\partial}{\partial y^n}\right)_{p_0}$$

de l'espace $\mathrm{T}_{p_0}\mathscr{X}_{\mathbb{R}}$ *correspondant à la carte* $(U, h_{\mathbb{R}})$.

Dans la suite nous n'étudierons plus les variétés $\mathbb{C}$-analytiques. L'extension — si elle est possible — des résultats prouvés pour les variétés différentiables aux variétés $\mathbb{C}$-analytiques sera ainsi laissée à l'entière initiative du lecteur.

* * *

Soient $\mathscr{X}$ et $\mathscr{Y}$ deux variétés différentiables (de dimensions respectives n et m) et soit $f: \mathscr{X} \to \mathscr{Y}$ une application différentiable quelconque.

Soient par ailleurs p un point arbitraire de $\mathscr{X}$, $q = f(p)$ son image dans $\mathscr{Y}$ par f, et

$$(U, h) = (U, x^1, \ldots, x^n) \quad \text{et} \quad (V, k) = (V, y^1, \ldots, y^m)$$

des cartes de $\mathscr{X}$ et $\mathscr{Y}$ respectivement, telles que $p \in U$ et $fU \subset V$.

On sait alors (cf. leçon 7) que l'application f se représente dans les cartes (U, h) et (V, k) par des formules de la forme

$$y^j = f^j(x^1, \ldots, x^n), \quad j = 1, \ldots, m,$$

où f^j sont des fonctions différentiables.

La $(n \times m)$-matrice

$$\left\| \left(\frac{\partial f^j}{\partial x^i}\right)_p \right\|, \quad i = 1, \ldots, n, \quad j = 1, \ldots, m, \tag{21}$$

dont les éléments sont les valeurs $\left(\frac{\partial f^j}{\partial x^i}\right)_p$ des dérivées partielles au point p (i.e. plus exactement au point $h(p) \in \mathbb{R}^n$), s'appelle *matrice jacobienne* de l'application f dans les cartes (U, h) et (V, k).

Cette matrice définit une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$ associant au vecteur $(a^1, \ldots, a^n) \in \mathbb{R}^n$ le vecteur $(b^1, \ldots b^m)$ de $\mathbb{R}^m$, où

$$b^j = \left(\frac{\partial f^j}{\partial x^i}\right)_p a^i, \quad j = 1, \ldots, m. \tag{22}$$

Les isomorphismes de coordonnées $A \mapsto A\ (U, h)$ et $A \mapsto A\ (V, k)$ permettent d'interpréter cette application comme une application linéaire de $\mathrm{T}_p\mathscr{X}$ dans $\mathrm{T}_q\mathscr{Y}$.

Définition 2. L'application de $\mathrm{T}_p\mathscr{X}$ dans $\mathrm{T}_q\mathscr{Y}$ qui vient d'être construite s'appelle *différentielle de l'application différentiable* f en p. On la désignera par $(\mathrm{d}f)_p$ ou simplement par $\mathrm{d}f_p$.

Cette construction a été effectuée à la leçon 3 pour les surfaces élémentaires.

Donc, si $A = (a^1, \ldots, a^n)$ dans la carte (U, h), alors $\mathrm{d}f_p(A) = (b^1, \ldots, b^m)$ dans la carte (V, k), où b^j, $j = 1, \ldots, m$, sont les nombres (22). Dans le langage de l'algèbre linéaire cela signifie que *l'application* $\mathrm{d}f_p$ *n'est autre que l'application linéaire* $\mathrm{T}_p\mathscr{X} \to \mathrm{T}_q\mathscr{Y}$

qui possède la matrice (21) *dans les bases*

$$\left(\frac{\partial}{\partial x^1}\right)_p, \ldots, \left(\frac{\partial}{\partial x^n}\right)_p \quad \textit{et} \quad \left(\frac{\partial}{\partial y^1}\right)_q, \ldots, \left(\frac{\partial}{\partial y^m}\right)_q.$$

Il faut bien sûr s'assurer que cette définition est intrinsèque, i.e. que l'application df_p est indépendante du choix des cartes (U, h) et (V, k).

Soient $(U', h') = (U', x^{1'}, \ldots, x^{n'})$ et $(V', k') = (V', y^{1'}, \ldots, y^{m'})$ d'autres cartes des variétés $\mathscr{X}$ et $\mathscr{Y}$, telles que $p \in U'$ et $q \in V'$, et soit $d'f_p$ l'application df_p construite à l'aide de ces cartes. Dans ce cas, si $A = (a^{1'}, \ldots, a^{n'})$ dans (U', h'), alors $d'f_p(A) = (b^{1'}, \ldots, b^{m'})$, où $b^{j'} = \left(\frac{\partial f^{j'}}{\partial x^{i'}}\right)_p a^{i'}$ et $\left(\frac{\partial f^{j'}}{\partial x^{i'}}\right)_p$ sont les valeurs en p des dérivées partielles des fonctions $f^{j'}$ par rapport à $x^{i'}$ exprimant l'application f dans les cartes (U', h') et (V', k').

Soient par ailleurs

$$x^i = x^i(x^{1'}, \ldots, x^{n'}) \quad \text{et} \quad y^{j'} = y^{j'}(y^1, \ldots, y^m)$$

les expressions des coordonnées locales x^i et $y^{j'}$ en fonction des coordonnées $x^{i'}$ et y^j.

Pour alléger les formules on posera

$$\mathbf{x} = (x^1, \ldots, x^n) \quad (\text{resp. } \mathbf{x}' = (x^{1'}, \ldots, x^{n'})),$$

$$\mathbf{y} = (y^1, \ldots, y^m) \quad (\text{resp. } \mathbf{y}' = (y^{1'}, \ldots, y^{m'})).$$

On écrira donc $x^i(\mathbf{x}')$ pour $x^i(x^{1'}, \ldots, x^{n'})$ et $\mathbf{x}(\mathbf{x}')$ pour $(x^1(\mathbf{x}'), \ldots, x^n(\mathbf{x}'))$.

Dans ces notations, les fonctions f^j et $f^{j'}$ seront reliées par la formule

$$\mathbf{f}'(\mathbf{x}') = \mathbf{y}'(\mathbf{f}(\mathbf{x}(\mathbf{x}'))),$$

d'où il résulte en vertu de la règle de dérivation d'une fonction composée que

$$\left(\frac{\partial f^{j'}}{\partial x^{i'}}\right)_p = \left(\frac{\partial y^{j'}}{\partial y^j}\right)_q \left(\frac{\partial f^j}{\partial x^l}\right)_p \left(\frac{\partial x^l}{\partial x^{i'}}\right)_p.$$

Puisque

$$a^i = \left(\frac{\partial x^l}{\partial x^{i'}}\right)_p a^i \quad \text{et} \quad b^j = \left(\frac{\partial f^j}{\partial x^i}\right)_p a^i,$$

il s'ensuit

$$b^{j'} = \left(\frac{\partial f^{j'}}{\partial x^{i'}}\right)_p a^i = \left(\frac{\partial y^{j'}}{\partial y^j}\right)_q \left(\frac{\partial f^j}{\partial x^i}\right)_p \left(\frac{\partial x^i}{\partial x^{i'}}\right)_p a^{i'} =$$

$$= \left(\frac{\partial y^{j'}}{\partial y^j}\right)_q \left(\frac{\partial f^j}{\partial x^i}\right)_p a^i = \left(\frac{\partial y^{j'}}{\partial y^j}\right)_q b^j,$$

i.e. les nombres $b^{j'}$ et b^j sont les coordonnées (dans les cartes (V', k') et (V, k)) d'un même vecteur tangent. Donc, $\mathrm{d}'f_p(A) = \mathrm{d}f_p(A)$ pour tout vecteur $A \in \mathbf{T}_p\mathscr{X}$ et, par suite, $\mathrm{d}'f_p = \mathrm{d}f_p$. □

Comparer avec les raisonnements homologues de la leçon 3.

On remarquera que l'application $\mathrm{d}f_p$ dépend uniquement du comportement local de l'application f dans un voisinage du point p, i.e. *si pour des applications $f, g: \mathscr{X} \to \mathscr{Y}$ il existe un voisinage U du point p tel que $f = g$ sur U, alors* $\mathrm{d}f_p = \mathrm{d}g_p$.

Lorsque $\mathscr{X}$ et $\mathscr{Y}$ sont des sous-variétés ouvertes des espaces $\mathbb{R}^n$ et $\mathbb{R}^m$ (donc, les espaces $\mathbf{T}_p\mathscr{X}$ et $\mathbf{T}_q\mathscr{Y}$ s'identifient naturellement à ces espaces), l'application $\mathrm{d}f_p : \mathbf{T}_p\mathscr{X} \to \mathbf{T}_q\mathscr{Y}$ est confondue de toute évidence avec la partie principale linéaire $(\mathbb{R}^n \to \mathbb{R}^m)$ de l'application f (i.e. est sa différentielle au sens de l'analyse élémentaire).

* * *

Soient $f: \mathscr{X} \to \mathscr{Y}$ et $g: \mathscr{Y} \to \mathfrak{Z}$ des applications différentiables. Si $(U, h) = (U, x^1, \ldots, x^n)$, $(V, k) = (V, y^1, \ldots, y^m)$ et $(W, l) = (W, z^1, \ldots, z^s)$ sont des cartes des variétés $\mathscr{X}$, $\mathscr{Y}$ et $\mathfrak{Z}$, telles que $fU \subset V$ et $gV = W$, et si $\mathbf{y} = \mathbf{f}(\mathbf{x})$ et $\mathbf{z} = \mathbf{g}(\mathbf{y})$ sont les fonctions définissant dans ces cartes les applications f et g (on se sert des notations introduites ci-dessus), l'application différentiable $g \circ f: \mathscr{X} \to \mathfrak{Z}$ sera visiblement définie dans les cartes (U, h) et (W, l) par la fonction $\mathbf{z} = \mathbf{g}(\mathbf{f}(\mathbf{x}))$. Donc, en vertu de la formule de dérivation d'une fonction composée, pour tout point $p \in U$ on aura les égalités

$$\left(\frac{\partial z^k}{\partial x^i}\right)_p = \left(\frac{\partial z^k}{\partial y^j}\right)_q \left(\frac{\partial y^j}{\partial x^i}\right)_p, \quad q = f(p),$$

qui expriment que l'application linéaire $\mathrm{d}(g \circ f)_p$ est la composée des applications linéaires $\mathrm{d}f_p$ et $\mathrm{d}g_q$:

$$\mathrm{d}(g \circ f)_p = \mathrm{d}g_q \circ \mathrm{d}f_p. \tag{23}$$

Cette formule s'appelle r è g l e e n c h a î n e.

* * *

Les cas où $\mathscr{Y} = \mathbb{R}$ et $m = n$ présentent un intérêt spécial.

Lorsque $\mathscr{Y} = \mathbb{R}$ (donc, l'application f est une fonction différentiable sur $\mathscr{X}$), la différentielle $\mathrm{d}f_p$ s'appelle généralement *gradient* de la fonction f.

En vertu de l'identification (7), le gradient est une application linéaire $\mathbf{T}_p\mathscr{X} \to \mathbb{R}$, i.e. un covecteur de l'espace $\mathbf{T}_p\mathscr{X}$ (un vecteur de l'espace dual $\mathbf{T}_p^*\mathscr{X}$, qui à propos s'appelle généralement *espace cotangent* à la variété $\mathscr{X}$ au point p). Par définition, sur tout vecteur

$A \in \mathrm{T}_p\mathscr{X}$ le covecteur $\mathrm{d}f_p$ prend la valeur

$$\mathrm{d}f_p(A) = \left(\frac{\partial f}{\partial x^i}\right)_p a^i. \tag{24}$$

Cette valeur s'appelle *dérivée de la fonction f suivant le vecteur A* et se désigne par Af. Donc,

$$Af = \mathrm{d}f_p(A)$$

et

$$Af = \left(\frac{\partial f}{\partial x^i}\right) a^i \quad \text{si} \quad A = (a^1, \ldots, a^n) \quad \text{dans} \quad (U, x^1, \ldots, x^n).$$

En particulier, $\left(\frac{\partial}{\partial x^i}\right)_p f = \left(\frac{\partial f}{\partial x^i}\right)_p$ pour tout $i = 1, \ldots, n$, ce qui explique le choix des notations pour les vecteurs de la base (5).

La formule (24) exprime que dans la base de l'espace $\mathrm{T}_p^*\mathscr{X}$ duale de la base (5) de l'espace $\mathrm{T}_p\mathscr{X}$, le covecteur $\mathrm{d}f_p$ a pour coordonnées $\left(\frac{\partial f}{\partial x^1}\right)_p, \ldots, \left(\frac{\partial f}{\partial x^n}\right)_p$. Donc, *primo*, cette base est composée des covecteurs

$$\mathrm{d}x_p^1, \ldots, \mathrm{d}x_p^n$$

(il est clair que le covecteur $\mathrm{d}f_p$ est défini aussi pour les fonctions f définies seulement dans un voisinage du point p) et, *secundo*,

$$\mathrm{d}f_p = \left(\frac{\partial f}{\partial x^1}\right)_p \mathrm{d}x_p^1 + \ldots + \left(\frac{\partial f}{\partial x^n}\right)_p \mathrm{d}x_p^n = \left(\frac{\partial f}{\partial x^i}\right)_p \mathrm{d}x_p^i.$$

Remarque 1. Attirons l'attention sur le fait que *le gradient est un covecteur*. En analyse, l'interprétation du gradient d'une fonction différentiable sur $\mathbb{R}^n$ comme un vecteur (cf. leçon 24 plus bas) est le résultat d'une identification tacite des vecteurs et covecteurs au moyen de la structure euclidienne standard sur $\mathbb{R}^n$.

* * *

Supposons maintenant que $n = m$.

Définition 3. On dit qu'une application différentiable $f : \mathscr{X} \to \mathscr{Y}$ de variétés de même dimension est *étale en un point $p \in \mathscr{X}$* (ou est un *difféomorphisme local*) si elle applique difféomorphiquement un voisinage U de p sur un voisinage $V = fU$ du point $q = f(p)$.

Il est évident que tout difféomorphisme de $\mathscr{X}$ sur $\mathscr{Y}$ sera une application étale (en tout point $p \in \mathscr{X}$). Réciproquement, d'après le théorème de différentiabilité de la fonction réciproque, *toute application bijective étale $f : \mathscr{X} \to \mathscr{Y}$ est un difféomorphisme*, et de plus, comme le montrent des calculs simples (cf. remarque 1 de la leçon 6),

la condition de bijectivité est généralement nécessaire (i.e. elle ne résulte pas de l'étalement).

De la formule (12) (appliquée au difféomorphisme $f|_U : U \to V$ et au difféomorphisme réciproque $g : V \to U$) il s'ensuit immédiatement (puisque la différentielle de l'application identique est de toute évidence l'application identique) que *la différentielle* $\mathrm{d}f_p : \mathrm{T}_p\mathscr{X} \to \mathrm{T}_q\mathscr{Y}$ *d'une application* $f : \mathscr{X} \to \mathscr{Y}$ *étale en un point* p *est un isomorphisme* (une application linéaire inversible). Il s'avère que réciproquement, *si la différentielle* $\mathrm{d}f_p : \mathrm{T}_p\mathscr{X} \to \mathrm{T}_q\mathscr{Y}$ *d'une application différentiable* $f : \mathscr{X} \to \mathscr{Y}$ *en un point* $p \in \mathscr{X}$ *est un isomorphisme, l'application* f *est étale en* p. En effet, affirmer que l'application $\mathrm{d}f_p$ est un isomorphisme revient à affirmer que le déterminant de la matrice (9) est non nul (pour des cartes (U, h) et (V, k) données et $p \in U$, $fU \subset V$). Ce déterminant n'étant autre que le jacobien $D\varphi$ de l'application différentiable $\varphi = k \circ f \circ h^{-1} : h(U) \to f(V)$, on en déduit en vertu du théorème de l'application réciproque (cf. leçon 6) que l'application φ est un difféomorphisme sur un voisinage W' du point $h(p) \in h(U)$ dans $h(U)$. Donc, l'application f sera un difféomorphisme sur le voisinage $U' = h^{-1}W'$ du point p dans $\mathscr{X}$, donc sera étale en p. □

Nous avons ainsi prouvé la proposition suivante connue sous le nom de théorème des applications étales :

Proposition 1. *Une application différentiable* $f : \mathscr{X} \to \mathscr{Y}$ *est étale en un point* $p \in \mathscr{X}$ *si et seulement si sa différentielle*

$$\mathrm{d}f_p : \mathrm{T}_p\mathscr{X} \to \mathrm{T}_q\mathscr{Y}, \quad q = f(p),$$

est un isomorphisme. □

Si une application $f : \mathscr{X} \to \mathscr{Y}$ étale en un point est un difféomorphisme sur le support U d'une carte (U, h), alors le couple $(V, k) = (fU, h \circ (f|_U)^{-1})$ sera de toute évidence une carte dans $\mathscr{Y}$. Ceci étant, l'application $\varphi : h(U) \to k(V)$ associée à f sera l'application identique, i.e. dans les coordonnées locales correspondantes $x^1, \ldots, x^n$ et $y^1, \ldots, y^n$ l'application f sera définie par les formules

$$y^i = x^i, \quad i = 1, \ldots, n. \tag{25}$$

Donc, *si une application* $f : \mathscr{X} \to \mathscr{Y}$ *est étale en un point* $p \in \mathscr{X}$, *il existe alors dans les variétés* $\mathscr{X}$ *et* $\mathscr{Y}$ *des coordonnées locales* $x^1, \ldots, x^n$ *et* $y^1, \ldots, y^n$ *dans lesquelles l'application* f *s'exprime par les formules* (25) (i.e. est une application d'égalisation des coordonnées).

* * *

La proposition 1 n'est en fait qu'une autre formulation du théorème de l'application réciproque. Il est intéressant de noter que le dernier théorème admet une autre incarnation d'un principe différent.

Soit $(U, h) = (U, x^1, \ldots, x^n)$ une carte d'une variété différentiable $\mathcal{X}$, centrée en un point $p_0 \in \mathcal{X}$ (i.e. telle que $p_0 \in U$), et soient

$$x^{i'} = x^{i'}(x^1, \ldots, x^n), \quad i' = 1', \ldots, n', \tag{26}$$

des fonctions différentiables définies dans un voisinage $U' \subset U$ du point p_0.

On rappelle (cf. leçon 6) que le symbole $x^{i'}$ de la formule (26) a un double sens : à gauche il désigne une fonction sur U', à droite, une fonction sur un voisinage $h(U')$ du point $\mathbf{x}_0 = h(p_0)$ dans l'espace $\mathbb{R}^n$. De ce fait, le jacobien

$$\det \left\| \frac{\partial x^{i'}}{\partial x^i} \right\|, \quad \begin{array}{l} i = 1, \ldots, n, \\ i' = 1', \ldots, n', \end{array} \tag{27}$$

des fonctions (26) sera aussi traité soit comme une fonction sur U', soit comme une fonction sur $h(U')$. Celui des deux sens dans lequel il est compris ressortira à chaque fois du contexte.

Si les fonctions (26) sont des coordonnées locales dans le voisinage U', le jacobien (27) est différent de 0 en chaque point de U' et en particulier en p_0. La réciproque peut être énoncée sous la forme suivante :

Proposition 2. *Si le jacobien* (27) *est différent de* 0 *au point* p_0, *les fonctions* (26) *seront des coordonnées locales dans un voisinage* $U'' \subset U'$ *de* p_0.

D é m o n s t r a t i o n. Les fonctions (26) traitées comme des fonctions sur l'ensemble ouvert $h(U')$ de l'espace $\mathbb{R}^n$ définissent une application φ de l'ensemble $h(U')$ dans $\mathbb{R}^n$. Si leur jacobien (27) est différent de 0 au point $\mathbf{x}_0 = h(p_0)$ de cet ensemble, il existe d'après le théorème de l'application réciproque un voisinage V' de $\mathbf{x}_0$ (contenu dans l'ensemble ouvert $h(U')$) sur lequel l'application φ est un difféomorphisme sur un ensemble ouvert de $\mathbb{R}^n$. Le couple (U'', h''), où

$$U'' = h^{-1}V, \quad h'' = (\varphi|_{V''}) \circ (h|_{U''}),$$

sera alors une carte de $\mathcal{X}$ liée à la carte (U, h) par les fonctions de passage (26). Donc, les fonctions (26), traitées comme des fonctions sur U'', seront les coordonnées locales correspondant à l'application de coordonnées h''. □

La proposition 2 est connue sous le nom de t h é o r è m e d e c h a n g e m e n t d e s c o o r d o n n é e s l o c a l e s.

* * *

Revenons maintenant à une application différentiable quelconque $f : \mathcal{X} \to \mathcal{Y}$, où $\dim \mathcal{X} = n$ et $\dim \mathcal{Y} = m$.

Définition 4. On appelle *rang* d'une application différentiable $f : \mathcal{X} \to \mathcal{Y}$ en un point $p \in \mathcal{X}$ le rang r de l'application linéaire

$df_p : T_p\mathscr{X} \to T_q\mathscr{Y}$, $q = f(p)$, i.e. le rang de la matrice de Jacobi (21) de l'application f au point p.

Il est clair que

$$0 \leqslant r \leqslant \min(n, m).$$

Le rang d'une matrice ne pouvant qu'augmenter par une faible variation de ses éléments, le rang de l'application f en un point arbitraire d'un voisinage assez petit du point p sera supérieur à son rang au point p_0. Mais il peut très bien être strictement supérieur.

Définition 5. On dit qu'une application $f : \mathscr{X} \to \mathscr{Y}$ est *localement plane* en un point $p \in \mathscr{X}$ s'il existe un voisinage U de p sur lequel le rang de l'application f est constant (est égal au rang r en p).

Proposition 3. *Si une application $f : \mathscr{X} \to \mathscr{Y}$ est localement plane en un point $p \in \mathscr{X}$, il existe dans les variétés $\mathscr{X}$ et $\mathscr{Y}$ des cartes $(U, x^1, \ldots, x^n)$ et $(V, y^1, \ldots, y^m)$ telles que $p \in U$, $fU \subset V$, et l'application f s'exprime dans les coordonnées locales $x^1, \ldots, x^n$ et $y^1, \ldots, y^m$ par les formules*

$$y^j = \begin{cases} x^j & si\ j = 1, \ldots, r, \\ 0 & si\ j = r+1, \ldots, m. \end{cases} \tag{28}$$

On remarquera que les coordonnées $x^{r+1}, \ldots, x^n$ ne figurent pas dans les formules (28.)

De façon suggestive, la proposition 3 affirme qu'au voisinage d'un point p une application localement plane en p est construite comme la projection $\mathbb{R}^n \to \mathbb{R}^r \subset \mathbb{R}^m$ suivant les $n - r$ derniers axes de coordonnées sur le sous-espace $\mathbb{R}^r \subset \mathbb{R}^m$ des points dont les $m - r$ dernières coordonnées sont nulles.

La proposition 3 sera prouvée à la prochaine leçon.

LEÇON 13

Démonstration du théorème des applications localement planes. — Immersion et submersion. — Sous-variétés d'une variété différentiable. — Sous-espace tangent à une sous-variété. — Définition locale d'une sous-variété. — Unicité de la structure d'une sous-variété. — Cas des sous-variétés plongées. — Théorème de l'image réciproque d'une valeur régulière. — Solutions de systèmes d'équations. — Le groupe SL (n) traité comme une sous-variété.

Prouvons la proposition 3 de la leçon précédente.

Soient tout d'abord sur des variétés $\mathscr{X}$ et $\mathscr{Y}$ des cartes

$$(U, h) = (U, x^1, \ldots, x^n) \quad \text{et} \quad (V, k) = (V, y^1, \ldots, y^m)$$

telles que $p \in U$ et $fU \subset V$, et supposons que l'application f s'exprime dans ces cartes par les formules

$$y^j = f^j(x^1, \ldots, x^n), \quad j = 1, \ldots, m.$$

Par hypothèse, la matrice rectangulaire composée des éléments $\left(\frac{\partial f^j}{\partial x^i}\right)_p$ est de rang r. Quitte à renuméroter les coordonnées, on peut, sans perdre en généralité, admettre que le mineur

$$\det\left\|\left(\frac{\partial f^j}{\partial x^i}\right)_p\right\|, \quad i, j = 1, \ldots, r, \tag{1}$$

est différent de 0.

Sans perdre ceci de vue, considérons dans un voisinage U d'un point p des fonctions $x^{1'}, \ldots, x^{n'}$ définies par la formule

$$x^{i'} = \begin{cases} f^i(x^1, \ldots, x^n) & \text{si } i = 1, \ldots, r, \\ x^i & \text{si } i = r+1, \ldots, n. \end{cases} \tag{2}$$

Le jacobien $\det\left\|\frac{\partial x^{i'}}{\partial x^i}\right\|$ est visiblement égal au déterminant (1), donc est différent de 0. Par conséquent, en vertu de la proposition 2 de la leçon 12, les fonctions (2) sont des coordonnées locales dans un voisinage U' de p.

En prenant maintenant le voisinage U' pour U, et les coordonnées $x^{1'}, \ldots, x^{n'}$ pour $x^1, \ldots, x^n$, on obtient des cartes (U, h) et (V, k) dans lesquelles l'application f est exprimée par les formules

$$y^j = \begin{cases} x^j & \text{si } j = 1, \ldots, r, \\ f^j(x^1, \ldots, x^n) & \text{si } j = r+1, \ldots, n. \end{cases} \tag{3}$$

Donc, dans ces cartes, la matrice de Jacobi de l'application f s'écrit

$$\left\| \begin{array}{cccccc} 1 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & 0 & \dots & 0 \\ \frac{\partial f^{r+1}}{\partial x^1} & \dots & \frac{\partial f^{r+1}}{\partial x^r} & \frac{\partial f^{r+1}}{\partial x^{r+1}} & \dots & \frac{\partial f^{r+1}}{\partial x^n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial f^m}{\partial f^1} & \dots & \frac{\partial f^m}{\partial x^r} & \frac{\partial f^m}{\partial x^{r+1}} & \dots & \frac{\partial f^m}{\partial x^n} \end{array} \right\|.$$

Le rang de cette matrice étant par hypothèse égal à r non seulement en p mais aussi dans un voisinage de p (qui sans perte de généralité peut être supposé confondu avec U), les dérivées

$$\begin{array}{ccc} \frac{\partial f^{r+1}}{\partial x^{r+1}}, & \dots, & \frac{\partial f^{r+1}}{\partial x^n}, \\ \dots & \dots & \dots \\ \frac{\partial f^m}{\partial x^{r+1}}, & \dots, & \frac{\partial f^m}{\partial x^n} \end{array}$$

sont identiquement nulles dans U. Puisque de toute évidence on peut admettre que le voisinage U est connexe, on en déduit que dans U (ou plus exactement dans $h(U)$) les fonctions $f^{r+1}, \dots, f^m$ ne dépendent pas de $x^{r+1}, \dots, x^{(n)}$, i.e. leurs valeurs en tout point $(x^1, \dots, x^r, \dots, x^n) \in h(U)$ peuvent être écrites sous la forme

$$f^{r+1}(x^1, \dots, x^r), \dots, f^m(x^1, \dots, x^r).$$

En d'autres termes, ces fonctions peuvent être traitées comme des fonctions sur l'ensemble $\mathrm{pr}\, h(U) \subset \mathbb{R}^r$ des points $(x^1, \dots, x^r)$ de $\mathbb{R}^r$ tels que $(x^1, \dots, x^r, x^{r+1}, \dots, x^n) \in h(U)$ au moins pour un choix des nombres $x^{r+1}, \dots, x^n$ (géométriquement, l'ensemble $\mathrm{pr}\, h(U) \subset \mathbb{R}^n$ n'est autre que la projection de l'ensemble $h(U) \subset \mathbb{R}^n$ sur le sous-espace $\mathbb{R}^r \subset \mathbb{R}^n$).

D'autre part, des formules (3) il résulte immédiatement que la projection $\mathrm{pr}\, k(V) \subset \mathbb{R}^r$ de l'ensemble $k(V) \subset \mathbb{R}^m$, qui est définie de façon analogue, contient l'ensemble $\mathrm{pr}\, h(U)$. Donc, sur un voisinage du point $q = f(p)$ (plus exactement sur le voisinage $k^{-1}((\mathrm{pr}\ h(U)) \times \mathbb{R}^{m-r})$) sont définies les fonctions

$$f^{r+1}(y^1, \dots, y^r), \dots, f^m(y^1, \dots, y^r).$$

Posons

$$y^{j*} = \begin{cases} y^j & \text{si } j = 1, \dots, r, \\ y^j - f^j(y^1, \dots, y^r) & \text{si } j = r+1, \dots, m. \end{cases}$$

Puisque

$$\det\left\|\frac{\partial y^{j'}}{\partial y^j}\right\| = \left\|\begin{matrix} 1 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & 0 & \dots & 0 \\ & & & 1 & \dots & 0 \\ & ? & & \dots & \dots & \dots \\ & & & 0 & \dots & 1 \end{matrix}\right\| = 1,$$

où le signe ? représente des éléments sans intérêt pour nous, la proposition 2 de la leçon 12 nous dit que les fonctions $y^{1'}, \dots, y^{n'}$ sont des coordonnées locales sur un voisinage V' du point q. Ceci étant, dans les cartes $(U, x^1, \dots, x^n)$ et $(V', y^{1'}, \dots, y^{n'})$ l'application f sera définie par les formules

$$y^{j'} = \begin{cases} x^j & \text{si } j = 1, \dots, r, \\ 0 & \text{si } j = r+1, \dots, m. \end{cases}$$

Pour achever la démonstration, il reste à désigner V' par V, et $y^{1'}, \dots, y^{n'}$ par $y^1, \dots, y^n$. □

* * *

Définition 1. Une application différentiable $f : \mathscr{X} \to \mathscr{Y}$, où $\mathscr{X}$ et $\mathscr{Y}$ sont des variétés différentiables de dimension respective n et m, est une *immersion* en un point p de $\mathscr{X}$ si son rang est égal à n en ce point (ce qui bien sûr n'est possible que pour $n \leqslant m$), i.e. si l'application

$$\text{(4)} \qquad df_p : \mathrm{T}_p\mathscr{X} \to \mathrm{T}_q\mathscr{Y}, \quad q = f(p),$$

est un monomorphisme.

De façon analogue, une application $f : \mathscr{X} \to \mathscr{Y}$ est une *submersion* en un point p de $\mathscr{X}$ si son rang est égal à m en ce point (donc, $n \geqslant m$), i.e. si l'application (4) est un épimorphisme.

Donc, une application f est une immersion ou une submersion si son rang prend sa valeur maximale (pour n et m donnés). Pour cette raison, l'immersion et la submersion sont appelées *applications de rang maximal.*

En vertu de la proposition 1 de la leçon 12, une application est simultanément une immersion et une submersion (pour $n = m$) si et seulement si elle est étale.

Une application $f : \mathscr{X} \to \mathscr{Y}$ qui est une immersion (resp. submersion) en chaque point p de $\mathscr{X}$ s'appelle simplement *immersion* (resp. *submersion*).

Il est clair qu'étant des applications de rang maximal la submersion et l'immersion sont localement planes en p. Donc, en vertu

pe la proposition 3 de la leçon 7, *pour toute submersion* $f : \mathcal{X} \to \mathcal{Y}$ *en un point* p *de* $\mathcal{X}$ *il existe des cartes* $(U, x^1, \ldots, x^n)$ *et* $(V, y^1, \ldots, y^m)$ *sur les variétés* $\mathcal{X}$ *et* $\mathcal{Y}$, *telles que* $p \in U$, $fU \subset V$, *et l'application* f *se représente par rapport aux coordonnées locales* $x^1, \ldots, x^n$ *et* $y^1, \ldots, y^m$ *par les formules*

$$y^1 = x^1, \ldots, y^m = x^m. \tag{5}$$

De façon imagée, cela signifie que dans les coordonnées correspondantes toute submersion est une projection locale $\mathbb{R}^n \to \mathbb{R}^m$ associant à un point $(x^1, \ldots, x^m, \ldots, x^n)$ de $\mathbb{R}^n$ le point $(x^1, \ldots, x^m)$ de $\mathbb{R}^m$.

Pour l'immersion on intervertira les notations et on admettra que f est une application de $\mathcal{Y}$ dans $\mathcal{X}$. Alors *pour toute immersion* $f : \mathcal{Y} \to \mathcal{X}$ *en un point* $q \in \mathcal{Y}$ *il existe des cartes* $(V, y^1, \ldots, y^m)$ *et* $(U, x^1, \ldots, x^n)$ *sur les variétés* $\mathcal{Y}$ *et* $\mathcal{X}$, *telles que* $q \in V$, $fV \subset U$, *et l'application* f *se représente par rapport aux coordonnées locales* $y^1, \ldots, y^m$ *et* $x^1, \ldots, x^n$ *par les formules*

$$x^1 = y^1, \ldots, x^m = y^m, \quad x^{m+1} = 0, \ldots, x^n = 0. \tag{6}$$

De façon suggestive, ceci exprime que dans les coordonnées correspondantes toute immersion est une injection locale de $\mathbb{R}^m$ dans $\mathbb{R}^n$ associant à un point $(y^1, \ldots, y^m) \in \mathbb{R}^m$ le point $(y^1, \ldots, y^m, 0, \ldots, 0) \in \mathbb{R}^n$.

* * *

Définition 2. On dit qu'une variété différentiable $\mathcal{Y}$ est une *sous-variété* d'une variété différentiable $\mathcal{X}$ si elle est contenue dans $\mathcal{X}$ et si l'injection

$$\iota : \mathcal{Y} \to \mathcal{X}, \quad \iota(p) = p, \tag{7}$$

est une immersion en tout point p de $\mathcal{Y}$ (et en particulier est une application différentiable).

Vu que $\mathcal{Y} \subset \mathcal{X}$, sur $\mathcal{Y}$ est définie une topologie $\mathbb{T}_{\mathcal{Y}/\mathcal{X}}$ induite par la topologie $\mathbb{T}_{\mathcal{X}}$ de $\mathcal{X}$. D'une façon générale (cf. plus bas), cette topologie est d i f f é r e n t e de la topologie $\mathbb{T}_{\mathcal{Y}}$ de $\mathcal{Y}$. On peut seulement affirmer que $\mathbb{T}_{\mathcal{Y}/\mathcal{X}} \subset \mathbb{T}_{\mathcal{Y}}$ (cette inclusion équivaut à la continuité de l'application ι).

Définition 3. Lorsque ι est un homéomorphisme sur son image (i.e. si $\mathbb{T}_{\mathcal{Y}/\mathcal{X}} = \mathbb{T}_{\mathcal{Y}}$), la variété $\mathcal{Y}$ s'appelle *sous-variété plongée*. Dans le cas contraire, $\mathcal{Y}$ s'appelle *sous-variété immergée*. (Du reste, le dernier terme est aussi utilisé comme synonyme du terme « sous-variété » quand il s'agit de souligner que la sous-variété $\mathcal{Y}$ n'est généralement pas plongée.)

L'immersion (7) est de toute évidence une application injective. Réciproquement, soit $f : \mathcal{Y} \to \mathcal{X}$ une immersion qui est une application injective, et soit $f = \iota \circ f'$ sa représentation par la composition d'une application bijective $f' : \mathcal{Y} \to f(\mathcal{Y})$ et de l'injection $\iota : f(\mathcal{Y}) \to \mathcal{X}$. L'application f' étant bijective, on peut grâce à elle transporter la structure différentiable de $\mathcal{Y}$ sur $\mathcal{Y}' = f(\mathcal{Y})$. Alors $\mathcal{Y}'$ sera une variété différentiable, f' un difféomorphisme et $\iota = f \circ (f')^{-1}$ une immersion (en tant que composée d'une immersion et d'un difféomorphisme), i.e. $\mathcal{Y}'$ sera une sous-variété de $\mathcal{X}$.

Donc, les sous-variétés de $\mathcal{X}$ sont les images dans $\mathcal{X}$ d'immersions arbitraires $\mathcal{Y} \to \mathcal{X}$ qui sont des applications injectives. Ceci étant, les sous-variétés plongées sont les images des immersions qui sont des monéomorphismes (i.e. des homéomorphismes sur leurs images).

Comme exemples de sous-variétés plongées citons les arcs réguliers simples et les surfaces élémentaires. Les immersions correspondantes sont les paramétrisations. Ceci étant, la condition de régularité des paramétrisations exprime que la paramétrisation est une immersion.

* * *

Puisque pour toute sous-variété $\mathcal{Y} \subset \mathcal{X}$ l'injection (7) est une immersion, pour tout point $p \in \mathcal{Y}$ l'application

$$d\iota_p : \mathbf{T}_p\mathcal{Y} \to \mathbf{T}_p\mathcal{X}$$

est un isomorphisme de l'espace vectoriel $\mathbf{T}_p\mathcal{Y}$ sur le sous-espace vectoriel $\operatorname{Im} d\iota_p = (d\iota_p)(\mathbf{T}_p\mathcal{Y})$ de l'espace $\mathbf{T}_p\mathcal{X}$. Ce sous-espace s'appelle *espace tangent à la sous-variété $\mathcal{Y}$ en p*. Généralement on l'identifie à $\mathbf{T}_p\mathcal{Y}$ (à l'aide de l'isomorphisme $d\iota_p$).

La situation est spécifique lorsque $\mathcal{X}$ est un espace affine $\mathcal{A}$, et donc pour tout point $p \in \mathcal{X}$ (et en particulier pour tout point $p \in \mathcal{Y}$) l'espace vectoriel $\mathbf{T}_p\mathcal{X}$ s'identifie à l'espace vectoriel $\mathcal{V}$ auquel est attaché $\mathcal{A}$. Dans ce cas, il est d'usage d'identifier le sous-espace $\mathbf{T}_p\mathcal{Y}$ à la sous-variété linéaire $p + \mathbf{T}_p\mathcal{Y}$ de l'espace affine $\mathcal{A}$ (i.e., de façon imagée, d'admettre que ses vecteurs sont d'origine p). Comparer avec la définition 3 de la leçon 3.

Etant une immersion, l'application (7) peut être représentée en coordonnées locales par les formules (6). Ceci exprime que pour tout point p de la sous-variété $\mathcal{Y}$ il existe une carte $(U, x^1, \ldots, x^n)$, $p \in U$, sur la variété $\mathcal{X}$, telle que, *primo*, sur un ensemble $V \subset U \cap \mathcal{Y}$ (ouvert dans $\mathcal{Y}$) les restrictions

$$y^1 = x^1|_V, \ \ldots, \ y^m = x^m|_V$$

des m premières coordonnées $x^1, \ldots, x^m$ sont des coordonnées locales sur V et, *secundo*, qu'un point $q \in U$ appartient à V si et seule-

ment si

(8) $$x^{m+1}(q)=0, \ \ldots, \ x^n(q)=0.$$

Les coordonnées $x^1, \ldots, x^n$ seront dites *compatibles* avec la sous-variété $\mathcal{Y}$.

Les coordonnées $y^1, \ldots, y^m$ définissent dans $\mathrm{T}_p\mathcal{Y}$ la base

(9) $$\left(\frac{\partial}{\partial y^1}\right)_p, \ \ldots, \ \left(\frac{\partial}{\partial y^m}\right)_p,$$

et les coordonnées $x^1, \ldots, x^n$ dans $\mathrm{T}_p\mathcal{X}$ la base

(10) $$\left(\frac{\partial}{\partial x^1}\right)_p, \ \ldots, \ \left(\frac{\partial}{\partial x^n}\right)_p.$$

Ceci étant, puisque l'injection $\iota : \mathcal{Y} \to \mathcal{X}$ se représente en ces coordonnées par les fonctions $y^1 = x^1, \ldots, y^m = x^m$, sa différentielle $d\iota_p$ transformera la base (9) en les m premiers vecteurs de la base (10). Ceci exprime en vertu de l'identification de $\mathrm{T}_p\mathcal{Y}$ à $\mathrm{Im}\, d\iota_p$ que

$$\left(\frac{\partial}{\partial y^1}\right)_p = \left(\frac{\partial}{\partial x^1}\right)_p, \ \ldots, \ \left(\frac{\partial}{\partial y^m}\right)_p = \left(\frac{\partial}{\partial x^m}\right)_p,$$

i.e. que *les vecteurs* $\left(\frac{\partial}{\partial x^1}\right)_p, \ldots, \left(\frac{\partial}{\partial x^m}\right)_p$ *forment une base du sous-espace* $\mathrm{T}_p\mathcal{Y}$ *de* $\mathrm{T}_p\mathcal{X}$.

* * *

Si une sous-variété $\mathcal{Y}$ est plongée dans $\mathcal{X}$, sans perdre en généralité on peut admettre que

(11) $$V = U \cap \mathcal{Y}.$$

En effet, l'ensemble V étant ouvert dans $\mathcal{Y}$ et la topologie de $\mathcal{Y}$ induite par celle de $\mathcal{X}$, il existe dans $\mathcal{X}$ un ensemble ouvert W tel que $V = W \cap \mathcal{Y}$. En remplaçant U par $U \cap W$, on peut faire en sorte que, sans perte de généralité, l'égalité (11) soit réalisée. □

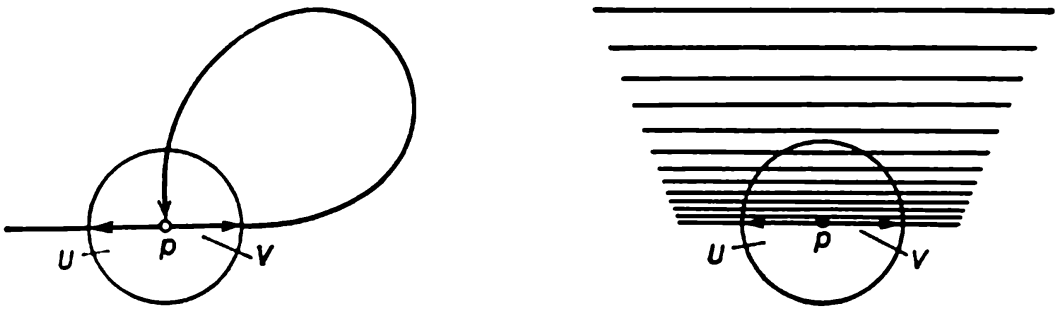
L'égalité (11) exprime qu'un point $q \in U$ appartient à $\mathcal{Y}$ si et seulement s'il vérifie les égalités (8). En d'autres termes, la variété $\mathcal{Y}$ est définie *localement* (i.e. dans le voisinage U) par les $n - m$ équations

(12) $$x^{m+1} = 0, \ \ldots, \ x^n = 0.$$

Dans le cas général, l'égalité (11) n'est généralement pas réalisée pour une sous-variété $\mathcal{Y}$ non plongée (immergée), et le lien entre V et $U \cap \mathcal{Y}$ est plus complexe.

L'ensemble V est fermé dans U (pour la topologie induite sur U par $\mathcal{T}_{\mathcal{Y}/\mathcal{X}}$) puisque les fonctions $x^{m+1}, \ldots, x^n$ sont continues et, donc, les conditions (8) définissent un sous-ensemble fermé dans U.

Par conséquent, V est fermé aussi dans $U \cap \mathcal{Y}$ (pour la topologie induite par $\mathbb{T}_{\mathcal{Y}/\mathcal{X}}$ et, par suite, pour celle induite par $\mathbb{T}_{\mathcal{Y}}$). D'autre part, l'ensemble V est par hypothèse ouvert dans $\mathcal{Y}$, donc dans $U \cap \mathcal{Y}$ (pour la topologie induite par $\mathbb{T}_{\mathcal{Y}}$). Donc l'ensemble V est ouvert-fermé dans $U \cap \mathcal{Y}$ (pour la topologie induite par $\mathbb{T}_{\mathcal{Y}}$). Donc, *si le voisinage de coordonnées V est connexe, c'est qu'il est la composante de l'ensemble $U \cap \mathcal{Y}$, qui contient le point p.*



Deux situations typiques dans lesquelles l'égalité $V = U \cap \mathcal{Y}$ est impossible

[On rappelle que tout point d'une variété différentiable possède un système fondamental de voisinages de coordonnées connexes — par exemple de voisinages homéomorphes aux boules ouvertes d'un espace euclidien de dimension respective.

* * *

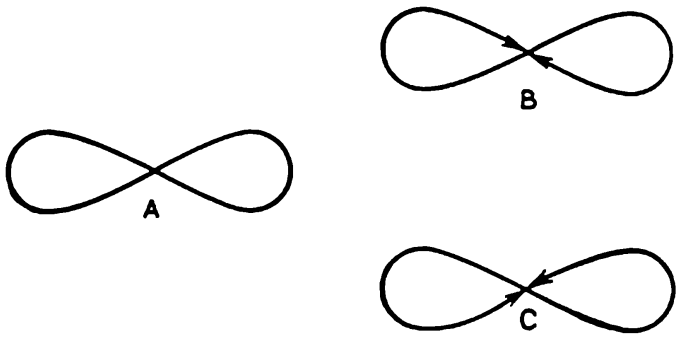
Supposons maintenant que sur un sous-ensemble $\mathcal{Y}$ d'une variété différentiable $\mathcal{X}$ sont données deux structures différentiables pour lesquelles $\mathcal{Y}$ est une sous-variété de $\mathcal{X}$ et qui définissent sur $\mathcal{Y}$ *la même topologie*. Alors les cartes de la forme $(V, y^1, \ldots, y^m)$ possédant les propriétés précitées seront les mêmes pour ces deux structures (puisque les fonctions $y^1, \ldots, y^m$ sont les restrictions des coordonnées locales $x^1, \ldots, x^m$, et les ensembles V, les composantes des ensembles $U \cap \mathcal{Y}$). D'autre part, ces cartes forment visiblement des sous-atlas d'un atlas maximal pour chacune de ces structures. Donc, ces deux structures sont confondues.

Donc, *pour une topologie donnée sur un sous-ensemble $\mathcal{Y}$ d'une variété différentiable $\mathcal{X}$, il existe sur $\mathcal{Y}$ au plus une structure de variété différentiable pour laquelle $\mathcal{Y}$ est une sous-variété de la variété $\mathcal{X}$.*

Certes, en faisant varier la topologie, on peut munir $\mathcal{Y}$ de plusieurs structures différentes de sous-variété. Par exemple, tout sous-ensemble $\mathcal{Y} \subset \mathcal{X}$ peut être muni d'une topologie discrète et transformé ainsi en sous-variété 0-dimensionnelle.

On obtient un exemple plus intéressant en envisageant sur le plan l'ensemble de la figure A (le « huit »). Cet ensemble ne peut être

une sous-variété du plan pour la topologie induite à cause du point singulier présenté au centre. Mais pour la topologie plus faible représentée schématiquement sur la figure B, il sera une variété immergée difféomorphe à la droite $\mathbb{R}$. Il sera aussi une sous-variété immergée difféomorphe à $\mathbb{R}$ pour une autre topologie, celle repré-



sentée schématiquement sur la figure C. Les injections correspondantes seront des courbes régulières non équivalentes sans points doubles ayant le même support.

* * *

S'agissant de la topologie induite sur $\mathcal{Y}$ par celle de la variété $\mathcal{X}$, on obtient en particulier que *sur un sous-ensemble $\mathcal{Y}$ d'une variété différentiable $\mathcal{X}$, il peut exister au plus une structure de variété différentiable pour laquelle $\mathcal{Y}$ est une sous-variété plongée.* On peut donc en toute légitimité dire que le sous-ensemble $\mathcal{Y}$ est une sous-variété plongée.

Pour dire si un sous-ensemble donné $\mathcal{Y}$ est une sous-variété plongée il faut

1° Envisager tous les couples de la forme $(V, y^1, \ldots, y^m)$ où V est l'intersection $U \cap \mathcal{Y}$ de $\mathcal{Y}$ avec le support U d'une carte $(U, x^1, \ldots, x^n)$ de la variété $\mathcal{X}$, et $y^1, \ldots, y^m$, les restrictions $x^1|_V, \ldots, x^m|_V$ des coordonnées locales $x^1, \ldots, x^m$ à V.

2° Choisir parmi eux les couples $(V, y^1, \ldots, y^m)$ qui sont des cartes sur $\mathcal{Y}$, i.e. les couples tels que les fonctions $y^1, \ldots, y^m$ définissent une application bijective de l'ensemble $V = U \cap \mathcal{Y}$ sur un ensemble ouvert $V \subset \mathbb{R}^m$.

Si les cartes choisies peuvent former un atlas sur $\mathcal{Y}$, alors $\mathcal{Y}$ sera une sous-variété plongée. En effet, il est clair que cet atlas munit $\mathcal{Y}$ d'une structure de sous-variété compatible avec la topologie induite. [A noter que dans le cas général l'atlas maximal de la variété $\mathcal{Y}$ ne peut contenir qu'une partie des cartes choisies.] □

* * *

Appliquons cette proposition générale à l'ensemble $\mathcal{Y} = f^{-1}(q_0)$, image réciproque d'un point $q_0 \in \mathfrak{Z}$ par une application différentiable $f: \mathcal{X} \to \mathfrak{Z}$.

Définition 4. On dit qu'un point $q_0 \in \mathfrak{Z}$ est une *valeur régulière* d'une application $f: \mathcal{X} \to \mathfrak{Z}$ si f est une submersion en tout point $p \in \mathcal{Y}$.

Proposition 1 (théorème de l'image réciproque d'une valeur régulière). *L'image réciproque $\mathcal{Y} = f^{-1}(q_0)$ d'une valeur régulière q_0 (si elle n'est pas vide) est une sous-variété plongée dans $\mathcal{X}$. Cette sous-variété est de dimension $n - r$, où $r = \dim \mathfrak{Z}$ (et $n = \dim \mathcal{X}$).*

D é m o n s t r a t i o n. Pour tout point $p \in \mathcal{Y}$ il existe sur $\mathcal{X}$ une carte $(U, x^1, \ldots, x^n) = (U, h)$ et sur $\mathfrak{Z}$ une carte $(W, z^1, \ldots, z^r)$ telles que $p \in U$, $fU \subset W$ et dans ces cartes l'application f est représentée par les formules

$$z^1 = x^{m+1}, \ \ldots, \ z^r = x^n, \tag{13}$$

où $m = n - r$. Ceci étant, on peut admettre sans perdre en généralité qu'au point q_0, les coordonnées $z^1, \ldots, z^r$ sont toutes nulles, donc qu'un point $u \in U$ appartient à $\mathcal{Y}$ (plus exactement à l'intersection $V = U \cap \mathcal{Y}$) si et seulement si

$$x^{m+1}(u) = 0, \ \ldots, \ x^n(u) = 0.$$

Il s'ensuit que si une application $k: V \to \mathbb{R}^m$ définie par la formule

$$k(u) = (y^1(u), \ldots, y^m(u)), \quad u \in U,$$

où

$$y^1 = x^1|_V, \ \ldots, \ y^m = x^m|_V,$$

est traitée, eu égard à l'injection canonique

$$\mathbb{R}^m \to \mathbb{R}^n, \quad (x^1, \ldots, x^m) \mapsto (x^1, \ldots, x^m, 0, \ldots, 0),$$

comme une application dans $\mathbb{R}^n$, elle ne sera autre que la restriction $h|_V$ à V d'une application $h: U \to h(U)$. Donc, l'ensemble $k(V)$ en tant qu'intersection $\mathbb{R}^m \cap h(U)$ sera ouvert dans $\mathbb{R}^m$, et l'application $k: V \to k(V)$ sera bijective. En d'autres termes, le couple $(V, k) = (V, y^1, \ldots, y^m)$ sera une carte sur $\mathcal{Y}$.

Soit maintenant $(U', x^{1'}, \ldots, x^{n'})$ une autre carte sur la variété $\mathcal{X}$ possédant les mêmes propriétés par rapport à l'application f et soit $(V', y^{1'}, \ldots, y^{m'})$ la carte correspondante sur $\mathcal{Y}$. Etant donné que sur l'intersection $U \cap U'$ les coordonnées $x^1, \ldots, x^n$ et $x^{1'}, \ldots, x^{n'}$ sont reliées par les formules

$$x^{i'} = x^{i'}(x^1, \ldots, x^n), \quad i' = 1', \ \ldots, \ n',$$

où

$$\det\left\|\frac{\partial x^{i'}}{\partial x^{i}}\right\| \neq 0 \quad \text{partout sur } U \cap U', \tag{14}$$

sur l'intersection $V \cap V' = U \cap U' \cap \mathcal{Y}$ les coordonnées $y^1, \ldots, y^m$ et $y^{1'}, \ldots, y^{m'}$ le seront par les formules

$$y^{j'} = x^{j'}(y^1, \ldots, y^m, 0, \ldots, 0), \quad j' = 1, \ldots, m. \tag{15}$$

D'autre part, l'ensemble $\mathcal{Y}$ étant défini dans U' par les équations

$$x^{m+1'} = 0, \ \ldots, \ x^{n'} = 0,$$

les relations suivantes

$$x^{k'}(y^1, \ldots, y^m, 0, \ldots, 0) = 0, \quad k' = m+1, \ \ldots, \ n,$$

seront vérifiées aussi dans $V \cap V' \neq \varnothing$. Il en résulte qu'en tout point de $V \cap V'$ les dérivées partielles $\frac{\partial x^{i'}}{\partial x^{l}}$, $l = 1, \ldots, m$, vérifieront les égalités

$$\frac{\partial x^{i'}}{\partial x^{l}} = \begin{cases} \dfrac{\partial y^{i'}}{\partial y^{l}} & \text{si } i' = 1, \ \ldots, \ m, \\ 0 & \text{si } i' = m+1, \ \ldots, \ n, \end{cases}$$

donc, le jacobien (14) se divisera par le jacobien

$$\det\left\|\frac{\partial y^{i'}}{\partial y^{l}}\right\|, \quad \begin{matrix} i' = 1', \ \ldots, \ m', \\ l = 1, \ \ldots, \ m. \end{matrix} \tag{16}$$

Donc, le jacobien (16) est différent de 0 et, par suite, les formules (15) définissent un difféomorphisme entre les ensembles respectifs.

Ceci prouve que deux cartes quelconques de la forme $(V, y^1, \ldots, y^m)$ sont compatibles sur $\mathcal{Y}$. Elles forment donc un atlas puisqu'elles recouvrent $\mathcal{Y}$, et, par suite, $\mathcal{Y}$ est une variété plongée. □

Des formules (13) qui définissent l'application f en coordonnées locales, il s'ensuit immédiatement que la différentielle

$$\mathrm{d}f_p : \mathrm{T}_p\mathcal{X} \to \mathrm{T}_{q_0}\mathcal{Z}$$

de cette application en $p \in \mathcal{Y}$ agit sur les vecteurs de base de l'espace $\mathrm{T}_p\mathcal{X}$ à l'aide des formules

$$\mathrm{d}f_p\left(\frac{\partial}{\partial x^{l}}\right)_p = \begin{cases} 0 & \text{si } i = 1, \ \ldots, \ m, \\ \left(\dfrac{\partial}{\partial z^{i-m}}\right)_{q_0} & \text{si } i = m+1, \ \ldots, \ n. \end{cases}$$

Donc, *les vecteurs*

$$\left(\frac{\partial}{\partial x^{1}}\right)_p, \ \ldots, \ \left(\frac{\partial}{\partial x^{m}}\right)_p \tag{17}$$

engendrent le noyau Ker df_p *de l'application* df_p, *et les vecteurs*

$$\left(\frac{\partial}{\partial x^{m+1}}\right)_p, \ldots, \left(\frac{\partial}{\partial x^n}\right)_p \tag{18}$$

un sous-espace isomorphe à l'espace tangent $\mathbf{T}_{q_0}\mathfrak{Z}$.

Donc,

$$\mathbf{T}_p\mathscr{X} = \operatorname{Ker} df_p \oplus \hat{\mathbf{T}}_{q_0}\mathfrak{Z},$$

où $\hat{\mathbf{T}}_{q_0}\mathfrak{Z}$ est le sous-espace engendré par les vecteurs (18).

D'autre part, les vecteurs (17) engendrent un sous-espace $\mathbf{T}_p\mathscr{Y}$ de l'espace $\mathbf{T}_p\mathscr{X}$ puisque $y^1 = x^1|_V, \ldots, y^m = x^m|_V$. Donc, *pour tout point* $p \in \mathscr{Y}$, *le noyau* Ker df_p *de l'application* df_p *est confondu avec l'espace tangent* $\mathbf{T}_p\mathscr{Y}$:

$$\mathbf{T}_p\mathscr{Y} = \operatorname{Ker} df_p. \tag{19}$$

et, par conséquent,

$$\mathbf{T}_p\mathscr{X} = \mathbf{T}_p\mathscr{Y} \oplus \hat{\mathbf{T}}_{q_0}\mathfrak{Z} \tag{20}$$

A noter que cette décomposition a lieu pour tout point $p \in \mathscr{Y}$.

Remarque 1. A signaler que toute sous-variété plongée $\mathscr{Y} \subset \mathscr{X}$ n'est pas l'image réciproque d'une valeur régulière par une application $f: \mathscr{X} \to \mathfrak{Z}$ (cf. exercice 1 plus bas).

En vertu de la formule (20), pour qu'une sous-variété plongée $\mathscr{Y} \subset \mathscr{X}$ soit l'image réciproque d'une valeur régulière, il faut (et on démontre qu'il suffit) que pour tout point $p \in \mathscr{Y}$ l'on ait la décomposition

$$\mathbf{T}_p\mathscr{X} = \mathbf{T}_p\mathscr{Y} \oplus N_p,$$

où N_p est un sous-espace pour lequel est définie (*pour tous les* p) une application isomorphe $N_p \to \mathbb{R}^{n-m}$ dépendant différentiablement du point p.

* * *

Soient $f^1, \ldots, f^r$ des fonctions différentiables sur une variété différentiable $\mathscr{X}$, et $a^1, \ldots, a^r$ des nombres réels. Un point $p \in \mathscr{X}$ est par définition une *solution* du système d'équations

$$f^1 = a^1, \ldots, f^r = a^r \tag{21}$$

si $f^i(p) = a^i$ pour tout $i = 1, \ldots, r$.

Soit $\mathscr{Y}$ l'ensemble de toutes les solutions du système (21).

On dit que les équations (21) sont *fonctionnellement indépendantes* si pour tout point $p \in \mathscr{Y}$ les covecteurs

$$df_p^1, \ldots, df_p^r \in \mathbf{T}_p^*\mathscr{X}$$

sont linéairement indépendants, donc les équations

$$df_p^1 = 0, \ldots, df_p^r = 0 \tag{22}$$

définissent un sous-espace vectoriel de $\mathbf{T}_p\mathscr{X}$ de dimension $m = n - r$.

Proposition 2. *L'ensemble $\mathcal{Y}$ des solutions de tout système* (21) *d'équations fonctionnellement indépendantes est une sous-variété plongée dans $\mathcal{X}$ de dimension $m = n - r$.*

Pour tout point $p \in \mathcal{Y}$, le sous-espace $\mathrm{T}_p\mathcal{Y}$ est un sous-espace de l'espace des solutions du système d'équations linéaires (22).

Démonstration. Les fonctions $f^1, \ldots, f^r$ définissent (au moyen de la formule $f(p) = (f^1(p), \ldots, f^r(p))$ une application différentiable $f: \mathcal{X} \to \mathbb{R}^r$ pour laquelle l'ensemble $\mathcal{Y}$ est l'image réciproque $f^{-1}(\mathbf{a})$ d'un point $\mathbf{a} = (a^1, \ldots, a^r) \in \mathbb{R}^r$. Ceci étant, l'indépendance fonctionnelle des équations (21) exprime de toute évidence que le point $\mathbf{a}$ est une valeur régulière de l'application f. Donc, la première assertion de la proposition 2 n'est qu'une reformulation de la proposition 1 pour le cas où $\mathfrak{Z} = \mathbb{R}^r$.

De façon analogue, l'espace des solutions des équations (22) n'étant autre que le noyau de l'application $\mathrm{d}f_p$, la deuxième assertion est une reformulation de l'égalité (19). □

Corollaire 1. *Soit f une fonction différentiable sur une variété différentiable $\mathcal{X}$ et soit $\mathcal{Y}$ l'ensemble de tous les points $p \in \mathcal{X}$ tels que $f(p) = a$, où $a \in \mathbb{R}$ est un nombre fixe. Si $\mathrm{d}f_p \neq 0$ pour tout point $p \in \mathcal{Y}$, alors $\mathcal{Y}$ est une sous-variété plongée de dimension $n - 1$ dont l'espace tangent en tout point $p \in \mathcal{Y}$ est l'hyperplan $\mathrm{d}f_p = 0$.* □

Si $f = f(x, y)$ est une fonction définie sur $\mathbb{R}^2$, la condition $\mathrm{d}f_p \neq 0$ exprime que soit $\frac{\partial f}{\partial x} \neq 0$, soit $\frac{\partial f}{\partial y} \neq 0$. Donc, la condition que le sous-ensemble du plan défini par l'équation $f(x, y) = 0$ est une variété plongée à une dimension exprime qu'il ne possède pas de points singuliers (cf. leçon 1).

Remarque 2. Le théorème de Whitney de la leçon 1 montre qu'il est généralement impossible de se dédouaner de la condition $\mathrm{d}f_p \neq 0$ dans le corollaire 1. Cependant, cette condition n'est nullement nécessaire pour que l'ensemble $\mathcal{Y}$ soit une sous-variété plongée (et forcément de dimension $n - 1$). Par exemple, dans $\mathbb{R}^3$ l'équation

$$x^2 + y^2 = 0$$

définit une droite (une sous-variété plongée), tandis que la différentielle $\mathrm{d}f_p$ de la fonction $f = x^2 + y^2$ est identiquement nulle en tous les points de cette droite.

De plus, d'après la remarque 1, toute sous-variété ne peut être définie par un système d'équations fonctionnellement indépendantes.

Exercice 1. Montrer que sur le plan projectif $\mathbb{R}P^2$ il n'existe pas de fonction différentiable f dont l'ensemble des zéros soit la droite projective $\mathbb{R}P^1 \subset \mathbb{R}P^2$.

Néanmoins dans la pratique, le fait qu'un sous-ensemble d'une variété différentiable est une sous-variété plongée s'établit généralement à l'aide de la proposition 2 (ou de son corollaire 1).

* * *

Exemple 1. Le déterminant $\det A$ d'une matrice carrée $A = \| a_{ij} \|$ est un polynôme de a_{ij} de degré 1 en chaque variable. Il est aisé de voir que

$$\frac{\partial \det A}{\partial a_{ij}} = A_{ij}, \tag{23}$$

où A_{ij} est le cofacteur de a_{ij}. (Il suffit de remarquer que d'après la formule du développement d'un déterminant suivant les éléments d'une colonne, $\det A = A_{ij}a_{ij} + A'_{ij}$, où A'_{ij} est indépendant de a_{ij}.) Puisque pour $\det A = 1$ on a nécessairement $A_{ij} \neq 0$ pour au moins un élément a_{ij}, on en déduit que les conditions du corollaire 1 de la proposition 2 sont remplies pour l'équation

$$\det A = 1$$

(sur la variété $\mathrm{GL}(n)$). *Le groupe* $\mathrm{SL}(n)$ *est donc une sous-variété de dimension* $n^2 - 1$ *plongée dans le groupe* $\mathrm{GL}(n)$.

Exercice 2. Montrer que la structure différentiable du groupe $\mathrm{SL}(n)$ est la même qu'il soit traité comme une sous-variété de $\mathrm{GL}(n)$ ou comme un groupe de Lie de matrices (cf. leçon 5).

Le groupe $\mathrm{GL}(n)$ étant une sous-variété ouverte de l'espace vectoriel $\mathrm{Mat}_n(\mathbb{R})$, pour toute matrice $A \in \mathrm{GL}(n)$ l'espace tangent $\mathrm{T}_A\,\mathrm{GL}(n)$ s'identifie de façon naturelle à $\mathrm{Mat}_n(\mathbb{R})$. Cette identification associe au vecteur de base $\left(\frac{\partial}{\partial a_{ij}}\right)$ de l'espace $\mathrm{T}_A\mathrm{GL}(n)$ la matrice unité $E_{ij} \in \mathrm{Mat}_n(\mathbb{R})$ et donc au covecteur $(\mathrm{d}a_{ij})_A$, le covecteur $\mathrm{Mat}_n(\mathbb{R}) \to \mathbb{R}$ qui envoie la matrice $B = \| b_{ij} \|$ dans son élément b_{ij}.

En vertu de la formule (23) on en déduit que la différentielle $\mathrm{d}f_A$ de la fonction $f : A \mapsto \det A$, traitée comme un covecteur de l'espace $\mathrm{Mat}_n(\mathbb{R})$, opère suivant la formule

$$\mathrm{d}f_A(B) = \sum_{i,j=1}^{n} A_{ij}b_{ij}, \qquad B = \| b_{ij} \| \in \mathrm{Mat}_n(\mathbb{R}),$$

où A_{ij} sont les cofacteurs des éléments de la matrice A. En particulier, pour $A = E$ (lorsque $A_{ij} = \delta_{ij}$) on trouve que

$$\mathrm{d}f_E(B) = \sum_{i=1}^{n} b_{ii} = \mathrm{Tr}\, B,$$

et donc que $\mathrm{Ker}\, \mathrm{d}f_E = \mathfrak{sl}(n)$.

Puisque $\mathrm{Ker}\, \mathrm{d}f_E = \mathrm{T}_E\mathrm{SL}(n)$, ceci prouve que *l'espace tangent au sous-groupe* $\mathrm{SL}(n)$ *au point* E *est son algèbre de Lie* $\mathfrak{sl}(n)$.

Remarque 3. Dans le prochain ouvrage (cf. Géométrie différentielle, leçon 13) on montrera que tout groupe de Lie de matrices est une sous-variété plongée dans la variété $\mathrm{GL}(n)$ dont l'espace tangent en E est son algèbre de Lie $\mathfrak{g}$.

LEÇON 14

Théorème de plongement. — Retour sur les ensembles compacts. — Fonctions d'Urysohn. — Démonstration du théorème de plongement. — Variétés satisfaisant au deuxième axiome de dénombrabilité. — Ensembles clairsemés et ensembles maigres. — Ensembles négligeables.

La classe de variétés la plus élémentaire et la plus suggestive est la classe des variétés difféomorphes pour un $N > 0$ aux sous-variétés plongées dans l'espace $\mathbb{R}^N$, et qui pour alléger les énoncés seront dites *plongeables* dans $\mathbb{R}^N$.

Il est clair que si un espace topologique est séparé et satisfait au deuxième axiome de dénombrabilité, il en sera de même de n'importe lequel de ses sous-espaces. Donc, *toute variété plongeable dans $\mathbb{R}^N$ est séparée et satisfait au deuxième axiome de dénombrabilité.* Il s'avère que cette condition nécessaire est suffisante aussi.

Théorème 1. *Pour toute variété séparée différentiable $\mathscr{X}$ satisfaisant au deuxième axiome de dénombrabilité, il existe un N tel que $\mathscr{X}$ soit plongeable dans $\mathbb{R}^N$.*

Que peut-on dire à propos de N?

Proposition 1. *Si une variété différentiable de dimension n est plongeable dans $\mathbb{R}^N$, où $N > 2n + 1$, elle le sera dans $\mathbb{R}^{N-1}$.*

Corollaire (théorème de plongement de Whitney). *Toute variété différentiable séparée $\mathscr{X}$ de dimension n satisfaisant au deuxième axiome de dénombrabilité est plongeable dans $\mathbb{R}^{2n+1}$.*

On prouvera le théorème 1 dans la seule hypothèse où *la variété $\mathscr{X}$ est compacte.* [On remarquera qu'une variété compacte satisfait (prouvez-le !) au deuxième axiome de dénombrabilité.] Le cas général implique des artifices supplémentaires d'ordre technique sur lesquels nous glisserons faute de temps.

Remarque 1. Le théorème 1 est valable pour les variétés d'une classe arbitraire C^r, $r \geqslant 1$, mais notre démonstration ne passera pas pour les variétés de classe C^1 et de classe C^ω ($\mathbb{R}$-analytiques), qui impliquent des raisonnements totalement différents et bien plus compliqués.

* * *

Commençons par quelques remarques simples sur les ensembles compacts.

Il est aisé de voir que *l'image continue d'un ensemble compact est compacte*, i.e. si $f: \mathscr{X} \to \mathscr{Y}$ est une application continue et $C \subset \mathscr{X}$ un sous-ensemble compact, l'ensemble $fC \subset \mathscr{Y}$ est compact. En effet, si des ensembles ouverts U_α recouvrent fC, les ensembles ouverts $f^{-1}U_\alpha$ recouvreront C, et si les ensembles $f^{-1}U_\alpha, \ldots, f^{-1}U_{\alpha_n}$ recouvrent C, les ensembles $U_{\alpha_1}, \ldots, U_{\alpha_n}$ recouvreront fC. □

En général, un sous-ensemble compact n'est pas obligatoirement fermé, mais si l'espace est séparé, il l'est nécessairement, autrement dit *tout sous-ensemble compact C d'un espace séparé $\mathscr{X}$ est fermé*. En effet, l'espace $\mathscr{X}$ étant séparé, pour tout point $c \in C$ et tout point $p \in \mathscr{X} \setminus C$ il existe des voisinages $U_p(c)$ et $V_c(p)$ de c et p respectivement tels que $U_p(c) \cap V_c(p) = \varnothing$. Puisque pour tout point $p \in \mathscr{X} \setminus C$ les voisinages $U_p(c)$, $c \in C$, forment un recouvrement de C, il existe en vertu de la compacité de C un système fini de points $c_1, \ldots, c_k \in C$ tel que

$$C \subset U_p(c_1) \cup \ldots \cup U_p(c_k).$$

Mais alors l'intersection

$$V = V_{c_1}(p) \cap \ldots \cap V_{c_k}(p) \tag{1}$$

sera un voisinage du point p ne rencontrant pas C, i.e. contenu dans $\mathscr{X} \setminus C$. Ceci prouve que tout point p de l'ensemble $\mathscr{X} \setminus C$ est un point intérieur, i.e. $\mathscr{X} \setminus C$ est ouvert. Donc, C est fermé. □

Il en découle que *si l'espace $\mathscr{X}$ est compact et l'espace $\mathscr{Y}$ séparé, toute application continue $f: \mathscr{X} \to \mathscr{Y}$ est fermée* (pour tout ensemble fermé $C \subset \mathscr{X}$ l'ensemble $fC \subset \mathscr{Y}$ est fermé). En effet, si l'ensemble C est fermé, il est compact puisque l'espace $\mathscr{X}$ l'est. Donc, l'ensemble fC est compact et, par suite, il est fermé puisque $\mathscr{Y}$ est séparé. □

Le cas particulier où l'application f est bijective est très important. En effet, puisque dans ce cas la fermeture de l'application f équivaut à la continuité de sa réciproque $f^{-1}: \mathscr{Y} \to \mathscr{X}$, on voit qu'*une application bijective continue d'un espace compact sur un espace séparé est un homéomorphisme*.

Si une application continue $f: \mathscr{X} \to \mathscr{Y}$ d'un espace compact $\mathscr{X}$ dans un espace séparé $\mathscr{Y}$ n'est qu'injective, c'est un monéomorphisme (un homéomorphisme sur son image $f(\mathscr{Y})$ qui dans ce cas est fermée dans $\mathscr{Y}$).

Donc, en particulier, *toute sous-variété compacte d'une variété séparée est une sous-variété plongée*.

* * *

Soient maintenant $\mathscr{X}$ une variété différentiable arbitraire, W et V des sous-ensembles ouverts de $\mathscr{X}$ tels que

$$\overline{W} \subset V. \tag{2}$$

Définition 1. On appelle *fonction d'Urysohn* du couple (V, W) une fonction différentiable $\varphi: \mathscr{X} \to \mathbb{R}$ telle que

a) $0 \leqslant \varphi(p) \leqslant 1$ pour tout point $p \in \mathscr{X}$;

b) $\varphi(p) = 1$ si et seulement si $p \in \overline{W}$;

c) si $p \in \mathscr{X} \setminus V$, alors $\varphi(p) = 0$.

On démontre que si la variété $\mathscr{X}$ est séparée et satisfait au deuxième axiome de dénombrabilité, *la fonction d'Urysohn existe pour tout couple* (V, W) *d'ensembles ouverts satisfaisant à la condition* (2). Mais nous n'aurons pas besoin de cette assertion en nous contentant de la démonstration du résultat plus faible suivant:

Proposition 2. *Si une variété différentiable* $\mathscr{X}$ *est séparée, il existe pour tout point* p_0 *de* $\mathscr{X}$ *et tout voisinage* U *de* p_0 *des ensembles ouverts* W *et* V *tels que*

$$p_0 \in W, \quad \overline{W} \subset V, \quad \overline{V} \subset U$$

et la fonction d'Urysohn existe pour le couple (V, W).

D é m o n s t r a t i o n. Il est clair que sans perdre en généralité on peut admettre que U est un voisinage de coordonnées du point p_0. Par définition ceci exprime qu'il existe un difféomorphisme h de l'ensemble U sur un ensemble ouvert $h(U) \subset \mathbb{R}^n$. Soit $\mathbf{x}_0 = h(p_0)$. Puisque $\mathbf{x}_0 \in h(U)$ et que l'ensemble $h(U)$ est ouvert, il existe un $r > 0$ tel que tout point $\mathbf{x} \in \mathbb{R}^n$ vérifiant $|\mathbf{x} - \mathbf{x}_0| < r$ (i.e. tout point de la boule $\mathbb{B}_r^n(\mathbf{x}_0)$) appartient à $h(U)$.

En remplaçant le difféomorphisme h par sa composée avec une similitude de l'espace $\mathbb{R}^n$, on peut donc, sans nuire à la généralité, admettre que $\mathbf{x}_0 = 0$ et $r = 3$, i.e. $\mathring{\mathbb{B}}_3^n \subset h(U)$, où (cf. leçon 1) $\mathring{\mathbb{B}}_3^n$ est la boule ouverte de $\mathbb{R}^n$ de rayon 3 centrée en 0.

Pour tout $r < 3$, l'ensemble $h^{-1}(\mathbb{B}_r^n)$ sera alors défini dans $\mathscr{X}$. Puisque la boule $\mathbb{B}_r^n$ est compacte et l'application h^{-1} est un homéomorphisme, cet ensemble sera aussi compact, donc fermé, car $\mathscr{X}$ est séparée par hypothèse.

D'autre part, l'ensemble $h^{-1}(\mathring{\mathbb{B}}_r^n)$ est ouvert (dans U, donc dans $\mathscr{X}$) et son adhérence $\overline{h^{-1}(\mathring{\mathbb{B}}_r^n)}$ contient de toute évidence l'ensemble $h^{-1}(\mathbb{B}_r^n)$. Donc,

$$\overline{h^{-1}(\mathring{\mathbb{B}}_r^n)} = h^{-1}(\mathbb{B}_r^n).$$

En posant

$$W = h^{-1}(\mathring{\mathbb{B}}_1^n) \quad \text{et} \quad V = h^{-1}(\mathring{\mathbb{B}}_2^n),$$

on en déduit immédiatement que

$$\overline{W} \subset V \quad \text{et} \quad \overline{V} \subset U.$$

Ceci étant, $p_0 \in W$ et

$$h(\overline{W}) = \mathbb{B}_1^n, \quad h(\overline{V}) = \mathbb{B}_2^n.$$

Rappelons-nous maintenant qu'en vertu du corollaire 2 du lemme 1 de la leçon 1, il existe sur l'espace $\mathbb{R}^n$ une fonction différentiable $f:\mathbb{R}^n \to \mathbb{R}$ telle que $0 \leqslant f \leqslant 1$ et

$$f(x) = \begin{cases} 1 & \text{si et seulement si } x \in \mathbf{B}_1^n, \\ 0 & \text{si et seulement si } x \notin \mathbf{B}_2^n. \end{cases}$$

Tout ceci nous permet de définir une fonction $\varphi : \mathscr{X} \to \mathbb{R}$ par la formule

$$\varphi(p) = \begin{cases} f(h(p)) & \text{si } p \in U, \\ 0 & \text{sinon.} \end{cases} \tag{3}$$

Il est clair que cette fonction possède les propriétés a), b) et c) de la définition 1. Il nous faut donc prouver seulement qu'elle est différentiable.

La fonction (3) est par construction différentiable sur U et sur $\mathscr{X} \setminus \bar{V}$ (elle est nulle sur le dernier ensemble). De plus, les ensembles ouverts U et $\mathscr{X} \setminus \bar{V}$ recouvrent $\mathscr{X}$. Il est clair par ailleurs que si une fonction $\varphi : \mathscr{X} \to \mathbb{R}$ est différentiable sur tout élément U_α d'un recouvrement ouvert $\{U_\alpha\}$ de $\mathscr{X}$. elle est différentiable sur $\mathscr{X}$ tout entière. Donc, en particulier, la fonction (3) est différentiable sur $\mathscr{X}$. □

* * *

Nous pouvons maintenant prouver le théorème 1.

Démonstration du théorème 1 (pour une variété $\mathscr{X}$ compacte). Choisissons pour chaque $p_0 \in \mathscr{X}$ un voisinage de coordonnées U et, en nous servant de la proposition 2, considérons l'ensemble W intervenant dans cette dernière. Les ensembles W construits pour tous les points $p_0 \in \mathscr{X}$ forment un recouvrement ouvert de $\mathscr{X}$. La variété $\mathscr{X}$ étant par hypothèse compacte, on peut extraire un sous-recouvrement fini de ce recouvrement. Désignons les ensembles W de ce sous-recouvrement par $W_1, \ldots, W_m$ et considérons les ensembles correspondants

$$V_1, \ldots, V_m, \quad U_1, \ldots, U_m,$$

les applications de coordonnées

$$h_1 : U_1 \to \mathbb{R}^n, \ \ldots, \ h_m : U_m \to \mathbb{R}^n$$

et les fonctions

$$\varphi_1 : \mathscr{X} \to \mathbb{R}, \ \ldots, \ \varphi_m : \mathscr{X} \to \mathbb{R}$$

de la proposition 2.

Ces ensembles, applications et fonctions possèdent les propriétés suivantes:

1° Pour tout $i = 1, \ldots, m$

$$\overline{W}_i \subset V_i, \quad \overline{V}_i \subset U_i.$$

2° Les familles

$$\{(W_1, h_1), \ldots, (W_m, h_m)\},$$
$$\{(V_1, h_1,), \ldots, (V_m, h_m)\},$$
$$\{(U_1, h_1,), \ldots, (U_m, h_m)\}$$

sont des atlas de la variété $\mathscr{X}$. (Il est évident que dans les deux premiers cas il est question des restrictions respectives des applications $h_1, \ldots, h_m$.)

3° Toute fonction φ_i, $i = 1, \ldots, m$, est la fonction d'Urysohn du couple (V_i, W_i).

En représentant les vecteurs de l'espace $\mathbb{R}^{n+1}$ sous la forme de couples $(\mathbf{x}, x)$, où $\mathbf{x} \in \mathbb{R}^n$ et $x \in \mathbb{R}$, on définit les applications

$$f_i : \mathscr{X} \to \mathbb{R}^{n+1}, \quad i = 1, \ldots, m,$$

par les formules

$$f_i(p) = \begin{cases} (\varphi_i(p)\, h_i(p), \varphi_i(p)) & \text{si } p \in U_i, \\ (0, 0) & \text{sinon.} \end{cases}$$

Il est clair que l'application f_i est différentiable sur U_i et nulle en dehors de $\overline{V}_i$. Elle est donc différentiable sur $\mathscr{X}$. D'autre part, $f_i(p) = (h_i(p), 1)$ pour tout point $p \in W_i$, d'où il résulte immédiatement (puisque h_i est un difféomorphisme) que l'application f_i est une immersion en tout point $p \in W_i$.

Considérons maintenant l'espace $\mathbb{R}^N$, où $N = m(n+1)$. En mettant les vecteurs de cet espace sous la forme $(\mathbf{x}_1, \ldots, \mathbf{x}_m)$, où $\mathbf{x}_1, \ldots, \mathbf{x}_m \in \mathbb{R}^{n+1}$, on définit une application $f : \mathscr{X} \to \mathbb{R}^N$ en posant pour tout point $p \in \mathscr{X}$

$$f(p) = (f_1(p), \ldots, f_m(p)).$$

Il est clair que l'application f est différentiable.

Soient p et q des points de $\mathscr{X}$ tels que $f(p) = f(q)$. Alors $f_i(p) = f_i(q)$ pour tout $i = 1, \ldots, m$ et en particulier $\varphi_i(p) = \varphi_i(q)$. D'autre part, puisque les cartes (W_i, h_i) forment un atlas, il existe un i_0 tel que $p \in W_{i_0}$ et donc $\varphi_{i_0}(p) = 1$ (et $f_{i_0}(p) = (h_{i_0}(p), 1)$). Par conséquent, $\varphi_{i_0}(q) = 1$, ce qui en vertu de la propriété b) des fonctions φ_i n'est possible que pour $q \in \overline{W}_i$. Donc, $f_{i_0}(q) = (h_{i_0}(q), 1)$ et, par suite, $h_{i_0}(p) = h_{i_0}(q)$. L'application h_{i_0} étant injective, ceci prouve que $p = q$. Donc, si $f(p) = f(q)$, alors $p = q$, i.e. *l'application f est injective*. La variété $\mathscr{X}$ étant compacte, l'application f est un monéomorphisme.

Considérons maintenant la différentielle df_p de l'application f en un point arbitraire $p \in \mathcal{X}$. En vertu de l'identification $T_{f(p)}\mathbb{R}^N = \mathbb{R}^N$, on peut admettre que

$$df_p : T_p\mathcal{X} \to \mathbb{R}^N.$$

Mais alors cette différentielle sera aussi composée des différentielles des applications f_i tout comme l'application f est composée des applications f_i, i.e. pour tout vecteur $A \in T_p\mathcal{X}$ on aura

$$(df)_pA = ((df_1)_pA, \ldots, (df_m)_pA).$$

Mais si, comme plus haut, $p \in W_{i_0}$ (donc, l'application f_{i_0} est une immersion en p, alors pour tout vecteur $A \in T_p\mathcal{X}$ non nul le vecteur $(df_{i_0})_pA \in \mathbb{R}^{n+1}$ est aussi non nul. Donc, il en est de même du vecteur $(df)_pA \in \mathbb{R}^N$. Ceci exprime que *l'application f est une immersion* (en tout point $p \in \mathcal{X}$).

Ainsi, l'application f est une immersion monéomorphe. Donc l'image $f\mathcal{X}$ est une sous-variété plongée dans l'espace $\mathbb{R}^N$, difféomorphe à la variété $\mathcal{X}$. □

* * *

Pour établir le théorème de plongement de Whitney (pour variétés compactes) il nous reste désormais à prouver seulement la proposition 1. Cette démonstration se fait en plusieurs temps.

Nous avons déjà remarqué que toute variété compacte (qui se recouvre par une famille finie de voisinages de coordonnées) satisfait au deuxième axiome de dénombrabilité. De façon plus générale, *toute variété recouvrable par un système dénombrable de voisinages de coordonnées satisfait au deuxième axiome de dénombrabilité.* En effet, étant homéomorphe à tout ouvert de $\mathbb{R}^n$, chaque voisinage de coordonnées satisfait à cet axiome. D'autre part, il est clair que tout espace admettant un recouvrement dénombrable dont les éléments satisfont (pour la topologie induite) au deuxième axiome de dénombrabilité. satisfait lui aussi à cet axiome. □

Réciproquement, il est aisé de voir que *dans toute variété différentiable $\mathcal{X}$ satisfaisant au deuxième axiome de dénombrabilité, il existe une base dénombrable composée de voisinages de coordonnées.* En effete, soit $\{O_\alpha\}$ une base dénombrable de $\mathcal{X}$. Pour tout point $p \in \mathcal{X}$ et tout voisinage de coordonnées U contenant le point p, il existe dans la base $\{O_\alpha\}$ un élément $O_{\alpha(p,U)}$ tel que $p \in O_{\alpha(p,U)} \subset U$. Soit O un sous-ensemble ouvert de $\mathcal{X}$. Choisissons pour tout point $p \in O$ un voisinage de coordonnées U_p tel que $p \in U_p \subset O$ et considérons l'ensemble

$$O_p = O_{\alpha(p,\, U_p)}.$$

On a $\bigcup_p O_p = O$, puisque $p \in O_p \subset U_p \subset O$ et p un point générique de O.

Donc, les ensembles de la forme $O_{\alpha(p,U)}$ (qui en vertu de l'inclusion $O_{\alpha(p,U)} \subset U$ sont des voisinages de coordonnées) forment une base dans $\mathcal{X}$. Pour achever cette démostration, il reste à remarquer que cette base est une partie de la base dénombrable $\{O_\alpha\}$, donc est elle-même une base dénombrable. □

* * *

Du cours d'analyse on sait que de nombreuses situations se simplifient considérablement si l'on se limite aux points génériques et que l'on néglige les ensembles « assez petits ». Il existe pour le moins deux approches différentes de la définition des ensembles « assez petits » de variétés différentiables : l'approche « topologique » et l'approche « métrique ».

La première approche repose sur la définition générale suivante :

Définition 2. On dit qu'un sous-ensemble d'un espace topologique $\mathcal{X}$ est *clairsemé* (ou *nulle part dense*) si son adhérence ne possède pas de points intérieurs.

On appelle *ensemble maigre* la réunion d'une famille finie ou dénombrable d'ensembles clairsemés.

Remarque 2. Le terme « ensemble maigre » a été introduit par Bourbaki. Naguère, les ensembles maigres s'appelaient « ensembles de première catégorie ». Cette dénomination n'est pas heureuse pour une foule de raisons et elle est tombée en désuétude.

A vrai dire un ensemble maigre peut être bien « gras » ; il peut par exemple être confondu avec l'espace $\mathcal{X}$ tout entier (tel le corps $\mathbb{Q}$ des rationnels et d'une façon générale tout espace topologique dénombrable ne possédant pas de points isolés).

Définition 3. On dit qu'un espace topologique $\mathcal{X}$ est un *espace de Baire* si aucun sous-ensemble maigre de $\mathcal{X}$ ne possède de points intérieurs.

Donc, dans un espace de Baire, les ensembles maigres sont littéralement « maigres ».

Les variétés différentiables sont-elles des espaces de Baire ? Il s'avère que oui (du moins pour les variétés séparées), mais la démonstration du théorème correspondant implique des préparatifs sur lesquels nous glisserons faute de temps. Nous préférerons donc l'approche « métrique » et nous prouverons sous forme d'exercices que les variétés séparées sont des espaces de Baire.

Définition 4. On dit qu'un espace topologique $\mathcal{X}$ est *régulier* si pour tout ouvert $U \subset \mathcal{X}$ et tout point $p \in U$ il existe un ouvert V tel que $p \in \bar{V}$ et $\bar{V} \subset U$.

Définition 5. On dit qu'un espace topologique $\mathcal{X}$ est *localement compact* si chacun de ses points p possède un voisinage O dont l'adhérence $\bar{O}$ est compacte.

Exercice 1. Montrer qu'*un espace $\mathscr{X}$ séparé localement compact est régulier.* [N o t a. Le sous-espace $\overline{O \cap U}$ est normal en vertu de la proposition 3 de la leçon 9. Il existe donc dans $\overline{O \cap U}$ un voisinage V de p, tel que $\overline{V} \subset O \cap U$.]

Exercice 2. Soit $\{A_n,\ n \geqslant 1\}$ une famille de sous-ensembles clairsemés d'un espace $\mathscr{X}$ séparé localement compact (donc, régulier) et soit U un voisinage d'un point $p \in \mathscr{X}$. Montrer qu'il existe une suite d'ensembles ouverts non vides U_n, $n \geqslant 0$, tels que:

a) l'ensemble $\overline{U_0}$ est compact et contenu dans U;

b) pour tout $n \geqslant 1$ on a

$$\overline{U}_n \subset U_{n-1}, \quad U_n \cap A_n = \varnothing.$$

[N o t a. L'ensemble A étant clairsemé, il existe dans l'ensemble U_{n-1} un ouvert non vide V_n tel que $V_n \cap A_{n-1} = \varnothing$ et, l'espace $\mathscr{X}$ étant régulier, il existe un ouvert non vide U_n tel que $\overline{U_n} \subset V_n$.]

Exercice 3. En déduire que *le voisinage U contient un point n'appartenant pas à la réunion A des ensembles A_n.* [N o t a. La suite $\{\overline{U}_n\}$ est une famille centrée d'ensembles fermés de l'espace compact $\overline{U_0}$, donc l'intersection des ensembles $\overline{U_n}$ n'est pas vide.]

L'assertion de l'exercice 3 exprime que *tout espace séparé localement compact est un espace de Baire.* Donc, toute variété séparée est un espace de Baire puisque de toute évidence elle est localement compacte.

* * *

On rappelle qu'un sous-ensemble A de l'espace euclidien $\mathbb{R}^n$ est un *ensemble de mesure nulle* (ou *ensemble négligeable*) si pour tout $\varepsilon > 0$ il existe une famille finie ou dénombrable de boules ouvertes de $\mathbb{R}^n$ recouvrant A, telle que la somme de leurs volumes (n-dimensionnels !) soit $< \varepsilon$.

Dans cette définition, les boules peuvent être remplacées par des parallélépipèdes (ou même des cubes) à arêtes parallèles aux axes de coordonnées, i.e. par des sous-ensembles de l'espace $\mathbb{R}^n$, dont les points $\mathbf{x} = (x^1, \ldots, x^n)$ sont caractérisés par des inégalités de la forme

$$a^i < x^i < b^i, \quad i = 1, \ldots, n.$$

Nous aurons besoin des trois propriétés suivantes des ensembles négligeables qui sont connues en analyse:

1° *La réunion $A = \bigcup_i A_i$ d'une famille finie ou dénombrable $\{A_i\}$ d'ensembles négligeables est un ensemble négligeable.*

2° *Pour toute application différentiable $f: U \to V$, où U et V sont des sous-ensembles ouverts de $\mathbb{R}^n$, et tout ensemble négligeable $A \subset U$ l'ensemble fA est un ensemble négligeable.*

3° *Aucun ensemble négligeable ne possède de points intérieurs.*

Pour prouver la propriété 1° il suffit de remarquer qu'en recouvrant chaque ensemble A_i, $i \geqslant 1$, par des boules dont la somme des volumes est $< \varepsilon/2^i$, on obtient un recouvrement de l'ensemble A par des boules dont la somme des volumes est $< \varepsilon$.

S'agissant de la propriété 2° on remarquera tout d'abord qu'en vertu de la propriété 1° il suffit de la prouver sous la condition subsidiaire qu'il existe un cube fermé $Q \subset U$ tel que A soit contenu dans $\mathring{Q}$. D'autre part, de la formule de Lagrange appliquée aux fonctions $f^1, \ldots, f^n$ définissant l'application f, il résulte immédiatement que pour tout couple de points $\mathbf{x}, \mathbf{y} \in Q$ on a la formule

$$|f(\mathbf{x}) - f(\mathbf{y})| \leqslant M \cdot |\mathbf{x} - \mathbf{y}|,$$

où M est une constante (le maximum des modules des dérivées premières des fonctions $f^1, \ldots, f^n$ dans le cube Q). Donc, toute boule de rayon r contenue dans le cube Q est envoyée par l'application f dans un ensemble contenu dans une boule de rayon Mr, donc de volume d'au plus M^n fois supérieur. Donc, en recouvrant l'ensemble A par des boules de volume total $<\varepsilon/M^n$, on obtient un recouvrement de l'ensemble fA par des boules de volume total $<\varepsilon$.

Pour prouver la propriété 3° il suffit de montrer que, si une famille finie $\{Q_1, \ldots, Q_m\}$ de cubes de l'espace $\mathbb{R}^n$, à arêtes parallèles aux axes de coordonnées, recouvre le cube Q, la somme des volumes des cubes $Q_1, \ldots, Q_m$ est supérieure au volume du cube Q (donc est minorée par une constante strictement positive ne dépendant pas du recouvrement). Soient $a_1, \ldots, a_m, a$ les longueurs respectives des arêtes des cubes $Q_1, \ldots, Q_m, Q$, et $N_1, \ldots, N_m, N$ le nombre de points de $\mathbb{R}^n$ à coordonnées entières contenus respectivement dans les cubes $Q_1, \ldots, Q_m, Q$. Les cubes $Q_1, \ldots, Q_m$ recouvrant Q, on a

$$N < N_1 + \ldots + N_m.$$

D'autre part il est clair que

$$[(a-1)^+]^n \leqslant N \leqslant (a+1)^n,$$

et de façon analogue

$$[(a_k-1)^+]^n \leqslant N_k \leqslant (a_k+1)^n$$

pour tout $k = 1, \ldots, m$, où

$$x^+ = \max(x, 0).$$

Donc,

$$[(a-1)^+]^n \leqslant \sum_{k=1}^{m} (a_k+1)^n.$$

En appliquant cette inégalité aux cubes homothétiques $\lambda Q_1, \ldots, \lambda Q_m$ et λQ, où $\lambda > 0$, on obtient pour λ assez grand l'inégalité

$$(\lambda a-1)^n \leqslant \sum_{k=1}^{m} (\lambda a_k+1)^n.$$

Donc,

$$\left(a-\frac{1}{\lambda}\right)^n \leqslant \sum_{k=1}^{m} \left(a_k+\frac{1}{\lambda}\right)^n,$$

d'où l'on déduit l'inégalité cherchée en faisant tendre λ vers $+\infty$:

$$a^n \leqslant \sum_{k=1}^{m} a_k^n.$$

Il est étonnant qu'un fait aussi évident implique une démonstration aussi ingénieuse !

Nous aurons besoin encore d'une propriété plus profonde des ensembles négligeables.

On appellera *coupe d'un sous-ensemble* $A \subset \mathbb{R}^n$ *suivant* $x_n = t$ le sous-ensemble A_t de l'espace $\mathbb{R}^{n-1}$, composé de points $\mathbf{x} = (x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1}$ tels que le point $(\mathbf{x}, t) = (x_1, \ldots, x_{n-1}, t)$ appartienne à A.

Théorème de Fubini. *Soit C un sous-ensemble compact de l'espace $\mathbb{R}^n$, tel que pour tout $t \in \mathbb{R}$ sa coupe C_t suivant $x_n = t$ est un ensemble négligeable (de l'espace $\mathbb{R}^{n-1}$). L'ensemble C est alors aussi un ensemble négligeable.*

Bien que ce théorème soit indiscutablement connu du cours d'analyse, nous allons le prouver intégralement afin que notre exposé soit complet.

On dira qu'un recouvrement de l'intervalle $[0, 1]$ par des intervalles ouverts est *admissible* si la réunion de ces intervalles est contenue dans l'intervalle $]-1, 2[$.

Lemme 1. *De tout recouvrement admissible de l'intervalle* $[0, 1]$ *par des intervalles ouverts, on peut extraire un sous-recouvrement fini dont la somme des longueurs des éléments est inférieure à* 6.

Démonstration. Montrons que tout sous-recouvrement minimal (i.e. un sous-recouvrement dont aucun élément ne peut être éliminé) possède la propriété requise.

Soient $]a_i, b_i[$, $1 \leqslant i \leqslant m$, les intervalles constituant ce sous-recouvrement minimal. Il est aisé de voir que $a_i \neq a_j$ pour $i \neq j$ (en effet, si $a_i = a_j$, alors pour $b_i \leqslant b_j$ l'intervalle $]a_i, b_i[$ est de trop, et pour $b_i \geqslant b_j$, l'intervalle $]a_i, b_j[$). Donc, quitte à renuméroter les intervalles, on peut admettre que $a_i < a_j$ pour $i < j$ (donc, en vertu de la minimalité, que $b_i < b_j$). Mais alors $a_i < a_{i+1} < b_i < a_{i+2}$ pour tout $i = 1, \ldots, m-2$, car pour $a_{i+1} \geqslant b_i$ il y aurait des trous dans le recouvrement et pour $b_i \geqslant a_{i+2}$ l'intervalle $]a_{i+1}, b_{i+1}[$ serait contenu, en vertu de la double inégalité $b_i < b_{i+1} < b_{i+2}$, dans la réunion des intervalles $]a_i, b_i[$ et $]a_{i+2}, b_{i+2}[$, donc serait superflu. Par conséquent, la somme des longueurs des intervalles du sous-recouvrement envisagé est justiciable de l'estimation

$$\begin{aligned}(b_1 - a_1) + (b_2 - a_2) + (b_3 - a_3) + \ldots + (b_{m-2} - a_{m-2}) + \\ + (b_{m-1} - a_{m-1}) + (b_m - a_m) \leqslant (a_3 - a_1) + (a_4 - a_2) + \\ + (a_5 - a_3) + \ldots + (a_m - a_{m-2}) + (b_{m-1} - a_{m-1}) + (b_m - a_m) = \\ = b_{m-1} + b_m - a_1 - a_2 < 2\,(b_m - a_1) \leqslant 6\end{aligned}$$

(car, en vertu de l'admissibilité, $-1 \leqslant a_1$ et $b_m \leqslant 2$). □

Démonstration du théorème de Fubini. Sans perdre en généralité on peut admettre de toute évidence que l'ensemble C est contenu dans la bande $0 \leqslant x_n \leqslant 1$. Par hypothèse, pour tout $\varepsilon > 0$ et tout $t \in [0, 1]$ il existe dans l'espace $\mathbb{R}^{n-1}$ des cubes ouverts $Q_t^{(i)}$ recouvrant la coupe C_t, dont le volume total est $< \varepsilon/6$. Soient Q_t leur réunion (qui contient la coupe C_t) et Q_t^* le sous-ensemble ouvert de l'espace $\mathbb{R}^n$ composé de tous les points $(\mathbf{x}, t)$, où $\mathbf{x} \in Q_t$ et $0 < t < 1$ (i.e. $Q_t^* = Q_t \times]0, 1[$). L'ensemble $C \setminus Q_t^*$ est compact en tant que sous-ensemble fermé d'un espace compact. D'autre part, la fonction $f(\mathbf{x}, x_n) = |x_n - t|$, $(\mathbf{x}, x_n) \in \mathbb{R}^n$, est continue et strictement positive sur l'ensemble $C \setminus Q_t^*$. Donc, il existe un nombre $\alpha > 0$ tel que

$$|x_n - t| > \alpha \quad \text{pour tout point } (\mathbf{x}, x_n) \in C \setminus Q_t^*. \tag{4}$$

Ceci étant, on peut, sans nuire à la généralité, admettre que $\alpha < 1$, i.e. que le recouvrement de l'intervalle $[0, 1]$ par les intervalles $I_t =]t - \alpha, t + \alpha[$ est admissible.

En vertu du lemme 1, on peut extraire de ce recouvrement un sous-recouvrement fini $\{I_{t_1}, \ldots, I_{t_m}\}$ composé d'intervalles dont la somme $2\alpha m$ des longueurs est < 6.

Soit L_{ij} un parallélépipède de l'espace $\mathbb{R}^n$, composé de points $(\mathbf{x}, \tau)$, $\mathbf{x} \in \mathbb{R}^{n-1}$, $\tau \in \mathbb{R}$, tels que $\mathbf{x} \in Q_{t_j}^{(i)}$ et $\tau \in I_{t_j}$ (i.e. $L_{ij} = Q_{t_j}^{(i)} \times I_{t_j}$). De la formule (4) il s'ensuit immédiatement que les parallélépipèdes recouvrent l'ensemble C. (En effet, soit $(\mathbf{x}, x_n) \in C$. Il existe un j tel que $x_n \in I_{t_j}$, i.e. tel que $|x_n - t_j| < \alpha$. Donc, $(\mathbf{x}, x_n) \in Q_{t_j}^*$ en vertu de la formule (4) et, par suite, $\mathbf{x} \in Q_{t_j}^{(i)}$ pour un certain i. Donc, $(\mathbf{x}, x_n) \in L_{ij}$.) D'autre part, le volume n-dimensionnel du parallélépipède L_{ij} est visiblement égal au volume $(n-1)$-dimensionnel du cube $Q_{t_j}^{(i)}$ multiplié par la longueur 2α de l'intervalle I_{t_j}. Donc, le volume total de tous les parallélépipèdes L_{ij} est inférieur à $m \cdot \frac{\varepsilon}{6} \cdot 2\alpha < \varepsilon$.

Donc, pour tout $\varepsilon > 0$, l'ensemble C peut être recouvert par des parallélépipèdes de volume total $< \varepsilon$. Cet ensemble est par conséquent négligeable. □

Les ensembles négligeables des variétés différentiables se définissent de façon naturelle.

Définition 6. On dit qu'un sous-ensemble A d'une variété différentiable à n dimensions $\mathcal{X}$ est un *ensemble négligeable* s'il existe dans $\mathcal{X}$ une famille finie ou dénombrable de cartes (U_i, h_i), telle que $A \subset \bigcup_i U_i$ et chaque ensemble $h_i(U_i \cap A)$ est négligeable dans $\mathbb{R}^n$.

Il est clair que *les propriétés* 1° *et* 2° *d'un ensemble négligeable dans* $\mathbb{R}^n$ *sont valables pour un ensemble négligeable dans toute variété différentiable* (dans la propriété 2°, par U et V il faut de toute évidence comprendre des variétés différentiables de même dimension). Il s'ensuit en particulier que *pour tout ensemble négligeable A et toute carte (U, h) de la variété $\mathcal{X}$ l'ensemble $h(U \cap A)$ est négligeable dans* $\mathbb{R}^n$. Si la variété $\mathcal{X}$ satisfait au deuxième axiome de dénombrabilité, on a réciproquement : *un sous-ensemble $A \subset \mathcal{X}$ est un ensemble négligeable si l'ensemble $h(U \cap A)$ l'est dans* $\mathbb{R}^n$ *pour toute carte (U, h) de la variété $\mathcal{X}$* (ou au moins pour toute carte d'une famille dénombrable de cartes dont les supports forment une base dans la variété $\mathcal{X}$).

Il est utile de comparer les ensembles négligeables avec les ensembles maigres. D'une façon générale, ces deux classes d'ensembles ne sont en aucune façon liées entre elles : il existe des ensembles maigres qui ne sont pas de mesure nulle (voire même de mesure totale, i.e. tels que leur complémentaire est de mesure nulle), et des ensembles négligeables qui ne sont pas maigres (voire même des ensembles négligeables dont le complémentaire est maigre).

Soit par exemple U_i, $1 \leqslant i < \infty$, la réunion d'un système dénombrable d'intervalles de la droite $\mathbb{R}$, dont les centres sont tous les points rationnels et dont la somme des longueurs est égale à 2^{-i}. L'ensemble $A = \bigcap_{i=1}^{\infty} U_i$ est alors négligeable et son complémentaire $\mathbb{R} \setminus A = \bigcup_{i=1}^{\infty} (\mathbb{R} \setminus U_i)$ est maigre, puisque $\mathbb{R} \setminus U_i$ est clairsemé pour tout i (il est fermé et ne possède pas de points intérieurs).

Les ensembles négligeables (dans $\mathbb{R}^n$, donc dans toute variété $\mathscr{X}$) ne possédant pas de points intérieurs, *tout ensemble négligeable fermé n'est nulle part dense*. Donc, tout ensemble négligeable qui se représente par la réunion d'une famille finie ou dénombrable d'ensembles fermés (qui sont nécessairement des ensembles négligeables) est un ensemble maigre. Ces ensembles ne possèdent pas de noms appropriés. Faute de mieux, on les appellera *ensembles négligeables maigres*.

LEÇON 15

Théorème de Sard. — Partie analytique de la démonstration du théorème de Sard. — Produit direct de variétés. — Variété de vecteurs tangents. — Démonstration du théorème de plongement de Whitney.

Soit $f : \mathcal{X} \to \mathcal{Y}$ une application différentiable d'une variété différentiable à n dimensions $\mathcal{X}$ dans une variété différentiable à m dimensions $\mathcal{Y}$.

Définition 1. On dit qu'un point $p \in \mathcal{X}$ est un *point critique* de l'application f si

$$\mathrm{d}f_p(\mathrm{T}_p\mathcal{X}) \neq \mathrm{T}_q\mathcal{Y}, \quad q = f(p),$$

i.e. si l'application

$$\mathrm{d}f_p : \mathrm{T}_p\mathcal{X} \to \mathrm{T}_q\mathcal{Y}$$

n'est pas épimorphe (l'application f n'est pas une submersion au point p). Un point $q \in \mathcal{Y}$ est une *valeur critique* de l'application f s'il existe un point critique $p \in \mathcal{X}$ tel que $q = f(p)$.

En comparant cette définition avec la définition 4 de la leçon 13, on voit aussitôt que *les valeurs critiques sont exactement les valeurs non régulières de l'application f.*

Remarquons que *pour $n < m$ tout point de la variété $\mathcal{X}$ est un point critique de l'application f.* Donc dans ce cas, les valeurs critiques sont les points de l'ensemble $f\mathcal{X}$, et les valeurs régulières, les points de son complémentaire $\mathcal{Y} \setminus f\mathcal{X}$. (Donc, *pour $n < m$ l'image réciproque de toute valeur régulière est vide.*)

Nous pouvons énoncer maintenant le théorème fondamental de cette leçon :

Théorème 1. *Si une variété $\mathcal{X}$ satisfait au deuxième axiome de dénombrabilité, l'ensemble $C(f)$ des valeurs critiques de toute application différentiable $f : \mathcal{X} \to \mathcal{Y}$ est négligeable ; si la variété $\mathcal{Y}$ est séparée, l'ensemble $C(f)$ est un ensemble négligeable maigre.*

Si la variété $\mathcal{X}$ est compacte (et la variété $\mathcal{Y}$ séparée), l'ensemble $C(f)$ est fermé et nulle part dense.

Ce théorème s'appelle t h é o r è m e d e S a r d, bien qu'il ait été démontré avant lui par Braun et indépendamment par Doubovitski.

Corollaire. *Si* $\dim \mathscr{X} < \dim \mathscr{Y}$, *aucune application différentiable* $f : \mathscr{X} \to \mathscr{Y}$ *n'est visiblement surjective* (*l'ensemble* $\mathscr{Y} \setminus f\mathscr{X}$ *n'est pas vide*). □

Nous aurons besoin uniquement de ce corollaire pour prouver le théorème de plongement de Whitney.

Signalons que dans le théorème 1 les variétés différentiables envisagées sont de classe C^∞ (ou C^ω). Du reste, un examen de la démonstration qui en sera donnée plus bas nous assure que ce théorème est aussi valable pour les variétés de classe C^r, où $r \geqslant n - m + 2$ pour $n \geqslant m$, et $r \geqslant 2$ pour $n \leqslant m$. [Bien plus, on peut diminuer r d'une unité encore par quelques artifices techniques. Des exemples montrent qu'il est toutefois impossible de réduire r davantage.]

Nous allons déduire le théorème de Sard de la proposition suivante :

Proposition 1. *Soit* $f : U \to \mathbb{R}^m$ *une application différentiable d'un ensemble ouvert* $U \subset \mathbb{R}^n$ *dans l'espace* $\mathbb{R}^m$, *et soit* K *un sous-ensemble compact de l'ensemble des points critiques de l'application* f. *L'ensemble* fK *est alors négligeable.*

Remarquons que, l'espace $\mathbb{R}^n$ étant séparé, l'ensemble fK est fermé, donc nulle part dense. D'autre part, l'ensemble des points critiques de l'application f est visiblement fermé et donc est la réunion d'une famille dénombrable de sous-ensembles compacts. Par conséquent, l'ensemble $C(f)$ des points critiques de l'application f est un ensemble négligeable maigre.

On voit donc que dans le cas particulier où $\mathscr{X} = U$, $\mathscr{Y} = \mathbb{R}^m$, le théorème de Sard est une conséquence immédiate de la proposition 1. Il se trouve que dans le cas général il se ramène aisément à cette proposition.

En effet, les points critiques de l'application f contenus dans tout sous-ensemble ouvert O de $\mathscr{X}$ ne sont visiblement autres que les points critiques de sa restriction $f|_O$ à O. Donc, pour toute base dénombrable $\{U_\alpha\}$ de la variété $\mathscr{X}$ on a

$$C(f) = \bigcup_\alpha C(f|_{U_\alpha}).$$

Ceci étant, si chaque U_α est le support d'une carte (U_α, f_α) (ce qui, on le sait, est toujours possible à réaliser) et si $f(U_\alpha) \subset V_\alpha$, où (V_α, k_α) est une carte de la variété $\mathscr{Y}$, alors $C(f|_{U_\alpha})$ est l'image par le difféomorphisme k_α^{-1} de l'ensemble $C(g_\alpha)$, où $g_\alpha = k_\alpha \circ (f|_{U_\alpha}) \circ h_\alpha^{-1} : U_\alpha \to \mathbb{R}^m$.

Donc, pour les mêmes raisons que plus haut, chaque ensemble

$$C(f|_{U_\alpha}) = k_\alpha^{-1} C(g_\alpha)$$

est la réunion d'une famille dénombrable d'ensembles négligeables compacts de la forme fK, où K est un ensemble compact composé de points critiques de l'application f.

Donc, la réunion d'une famille dénombrable de tels ensembles sera l'ensemble $C(f)$ tout entier. La réunion d'une famille dénombrable d'ensembles négligeables étant un ensemble négligeable, ceci prouve que l'ensemble $C(f)$ est un ensemble négligeable.

Si la variété $\mathcal{Y}$ est séparée, les ensembles fK sont tous fermés et, par conséquent, l'ensemble $C(f)$ est négligeable maigre.

Enfin, si de plus la variété $\mathcal{X}$ est compacte, l'ensemble de tous les points critiques de l'application f peut être décomposé en la réunion $K_1 \cup \ldots \cup K_N$ d'un nombre fini d'ensembles compacts contenus chacun dans un voisinage de coordonnées U_α. Dans ce cas donc $C(f) = fK_1 \cup \ldots \cup fK_N$, où, puisque $\mathcal{Y}$ est séparée, tous les ensembles fK_i sont fermés, donc nulle part denses. L'ensemble $C(f)$ est par conséquent fermé et nulle part dense. □

Il ne nous reste plus qu'à établir la proposition 1.

Remarquons tout d'abord que *la proposition* 1 *se prouve de façon triviale pour* $n < m$. En effet, si on admet que $\mathbb{R}^m = \mathbb{R}^n \times \mathbb{R}^{m-n}$, on peut envisager un ensemble ouvert $U \times \mathbb{R}^{m-n}$ de l'espace $\mathbb{R}^m$ et une application différentiable

$$F : U \times \mathbb{R}^{m-n} \to \mathbb{R}^m$$

qui est la composée de la projection sur le premier facteur $U \times \mathbb{R}^{m-n} \to U$ et d'une application $f : U \to \mathbb{R}^m$. Ceci étant, pour tout ensemble compact $K \subset U$ on aura

$$f(K) = F(K \times 0),$$

où $K \times 0$ est un sous-ensemble de l'espace $\mathbb{R}^m$, composé de points de la forme $(\mathbf{x}, 0)$, où $\mathbf{x} \in K$ et 0 est le vecteur nul de l'espace $\mathbb{R}^{m-n}$. Il est clair d'autre part que l'ensemble $K \times 0$ est un ensemble négligeable de l'espace $\mathbb{R}^m$ (puisqu'on peut le recouvrir par un système fini de parallélépipèdes de hauteur aussi petite que l'on veut). Donc, l'ensemble $F(K \times 0)$ est aussi un ensemble négligeable en vertu de la propriété 2° des ensembles négligeables de la leçon 14. □

Ceci prouve le théorème de Sard pour $n < m$, donc son corollaire formulé plus haut (qui, on le rappelle, ne nous servira qu'à prouver le théorème de Whitney).

* * *

Passons maintenant à la démonstration de la proposition 1 dans le cas général. Bien que cette proposition relève à proprement parler de l'analyse, nous allons la prouver soigneusement en dépit de sa lourdeur. (Le lecteur qui ne s'intéresse qu'au théorème de Whitney peut glisser sur cette démonstration.)

Soit $f : U \to \mathbb{R}^m$ une application définie par des fonctions $f^1, \ldots, f^m : U \to \mathbb{R}$ et soit C_k, où $k \geqslant 1$, l'ensemble de tous les points de U en lesquels s'annulent toutes les dérivées partielles d'ordre $\leqslant k$ des fonctions $f^1, \ldots, f^m$. Il est

clair que

$$C_0 \supset C_1 \supset \ldots \supset C_k \supset \ldots,$$

où C_0 est l'ensemble de tous les points critiques de l'application f $\left(\text{i.e. des points en lesquels le rang de la matrice } \left\| \frac{\partial f^j}{\partial x^i} \right\| \text{ est } < m\right)$.

Lemme 1. *Si $C_{s+1} = \varnothing$ ou $s > \frac{n}{m} - 1$, pour tout ensemble compact $K \subset C_s$ l'ensemble fK est négligeable.*

Ce lemme permet de démontrer la proposition 1 par une récurrence peu compliquée.

En effet, il existe un s tel que l'ensemble $f(K \cap C_s)$ soit négligeable (d'après le lemme, tout $s > \frac{n}{m} - 1$ convient). D'autre part, $K \cap C_0 = K$. Donc, pour prouver la proposition 1, il suffit de montrer que *si pour un $s \geqslant 1$ l'ensemble $f(K \cap C_s)$ est négligeable, il en est de même de l'ensemble $f(K \cap C_{s-1})$.*

Soit K_r, $r \geqslant 1$, un sous-ensemble de l'ensemble $K \cap C_{s-1}$, composé des points $\mathbf{x} \in K \cap C_{s-1}$ tels que $|\mathbf{x} - \mathbf{y}| \geqslant 1/r$ pour tout point $\mathbf{y} \in K \cap C_s$. L'ensemble K_s est compact en tant que sous-ensemble fermé de l'ensemble compact K.

Il est clair que

$$K \cap C_{s-1} = (K \cap C_s) \cup \bigcup_{r=1}^{\infty} K_r.$$

Il nous faut prouver seulement que *pour tout $r \geqslant 1$ l'ensemble fK_r est négligeable.*

Considérons à cet effet l'ensemble ouvert $U^* = U \setminus C_s$. Soit $f^* = f|_{U^*}$ et soient C_k^* les ensembles C_k construits pour la fonction f^*. Il est clair que $C_k^* = C_k \setminus C_s$ et donc $C_s^* = \varnothing$ et $K_r \subset C_{s-1}^*$. L'application f^* et l'ensemble compact K_r sont donc justiciables du lemme 1. Par conséquent, l'ensemble fK_r est bien négligeable. □

Il reste maintenant à prouver le lemme 1.

Démonstration du lemme 1. Ce lemme est composé de deux propositions à prémisses différentes mais avec la même conclusion. Considérons tout d'abord la première proposition (avec la prémisse $C_{s+1} = \varnothing$). L'ensemble K étant compact, pour prouver cette proposition il suffit de trouver pour tout point $\mathbf{x}_0 \in K$ un même voisinage $V \subset U$ tel que l'ensemble $f(K \cap V)$ soit négligeable. A cet effet, cherchons un voisinage V du point $\mathbf{x}_0$, tel que $\overline{V}$ soit compact et $f(C_s \cap \overline{V})$ soit négligeable. L'ensemble $f(K \cap V) \subset f(C_s \cap \overline{V})$ sera alors négligeable aussi.

Raisonnons par récurrence sur n. Cette assertion étant visiblement vérifiée pour $n = 0$, il nous faut prouver que si elle l'est pour $n - 1$, elle le sera pour n.

La condition $C_{s+1} = \varnothing$ exprime que la dérivée d'ordre $s + 1$ de l'une des fonctions $f^1, \ldots, f^m$ est différente de 0 en $\mathbf{x}_0$. Quitte à intervertir les coordonnées, on peut, sans perdre en généralité, admettre que cette dérivée est de la forme $\frac{\partial \varphi}{\partial x_n}$, où φ est une dérivée d'ordre s.

Soit g une application de U dans $\mathbb{R}^m$, définie par

$$g(\mathbf{x}, x_n) = (\mathbf{x}, \varphi(\mathbf{x}, x_n)), \qquad (\mathbf{x}, x_n) \in U.$$

Le jacobien de cette application est égal à $\frac{\partial \varphi}{\partial x_n}$, donc il est différent de 0 en $\mathbf{x}_0$. Il existe par conséquent un voisinage O' de $\mathbf{x}_0$ tel que l'application g est

un difféomorphisme de ce voisinage sur un ouvert O et, par suite, est définie l'application

$$h : O \xrightarrow{(g|_{O'})^{-1}} O' \xrightarrow{f|_{O'}} \mathbb{R}^m.$$

Soit V un voisinage du point $\mathbf{x}_0$ tel que $\overline{V}$ soit contenue dans O' et soit compacte. Montrons que *l'ensemble* $f(C_s \cap \overline{V})$ *est négligeable.*

On distinguera deux cas: $s = 0$ et $s > 0$. Commençons par le cas $s > 0$ qui est plus facile.

Pour $s > 0$ la fonction φ est nulle sur C_s en tant que dérivée d'ordre s de l'une des fonctions $f^1, \ldots, f^m$. Donc, l'ensemble $g(C_s)$ est contenu dans le sous-espace $\mathbb{R}^{n-1} \times \{0\}$ de l'espace $\mathbb{R}^n$, défini par l'équation $x_n = 0$ et, par suite,

$$f(C_s \cap \overline{V}) = h_0(g(C_s \cap \overline{V})),$$

où h_0 est la restriction de l'application h à $(\mathbb{R}^{n-1} \times \{0\}) \cap O$. Ceci prouve de toute évidence ce qu'on voulait, puisque l'application h_0 est en fait une application d'un ensemble ouvert de l'espace $\mathbb{R}^{n-1}$, et par suite, l'ensemble $h_0(g(C_s \cap \overline{V}))$ est négligeable d'après l'hypothèse de récurrence (appliquée à l'application h_0 et à l'ensemble compact $g(C_s \cap \overline{V})$).

Pour $s = 0$ le rôle de la dérivée φ incombe à l'une des fonctions $f^1, \ldots, f^m$ (qui est généralement non nulle sur C_0). Quitte à intervertir les coordonnées dans $\mathbb{R}^m$, on peut, sans perdre en généralité, admettre que $\varphi = f^m$.

Soit O_t la coupe suivant $x_n = t$ de l'ensemble O (l'ensemble de tous les points $\mathbf{x} \in \mathbb{R}^{n-1}$ tels que le point $(\mathbf{x}, t) \in \mathbb{R}^n$ appartienne à O).

Si $\mathbf{x} \in O_t$ et $g(\mathbf{x}, a) = (\mathbf{x}, t)$, alors par définition $f^m(\mathbf{x}, a) = \varphi(\mathbf{x}, a) = t$ et, par conséquent,

$$h(\mathbf{x}, t) = f(\mathbf{x}, a) = (f^1(\mathbf{x}, a), \ldots, f^{m-1}(\mathbf{x}, a), t).$$

Ceci exprime qu'en posant

$$h_t(\mathbf{x}) = (f^1(\mathbf{x}, a), \ldots, f^{m-1}(\mathbf{x}, a)),$$

où $\mathbf{x} \in O_t$ et $g(\mathbf{x}, a) = (\mathbf{x}, t)$ (i.e. $t = f^m(\mathbf{x}, a)$), on obtient une application différentiable

$$h_t : O_t \to \mathbb{R}^{m-1}$$

telle que

$$h(\mathbf{x}, t) = (h_t(\mathbf{x}), t) \tag{1}$$

pour tout point $(\mathbf{x}, t) \in O$.

De la formule (1) il s'ensuit que la matrice jacobienne J de l'application h est de la forme

$$\left\| \begin{array}{ccc|c} & & & * \\ & J_t & & \vdots \\ & & & * \\ \hline 0 & \cdots & 0 & 1 \end{array} \right\|,$$

où J_t est la matrice jacobienne de l'application h_t.

Donc, le rang de la matrice J au point $(\mathbf{x}, t)$ est $< m$ si et seulement si celui de la matrice J_t au point $\mathbf{x}$ est $< m - 1$. Mais par définition le rang de la matrice

J_t en $\mathbf{x} \in O_t$ est $< m - 1$ si et seulement si le point $\mathbf{x}$ est un point critique de l'application h_t, et par analogie, le rang de la matrice J en $(\mathbf{x}, t) \in O$ est $< m$ si et seulement si $\mathbf{x}$ est un point critique de l'application h (et, par suite, le point $(\mathbf{x}, a) = g^{-1}(\mathbf{x}, t) \in O'$ est un point critique de l'application f, i.e. appartient à l'ensemble $C_0 \cap O'$). Donc, la coupe $g(C_0 \cap O')_t$ de l'ensemble $g(C_0 \cap O')$ n'est autre que l'ensemble des points critiques de l'application h_t. Donc, d'après l'hypothèse de la récurrence (appliquée à h_t et à l'ensemble compact $g(C_0 \cap O')_t \cap g(\overline{V})_t)$), le sous-ensemble

$$h_t(g(C_0 \cap O')_t \cap g(\overline{V})_t)$$

de l'espace $\mathbb{R}^{n-1}$ est négligeable.

L'ensemble précédent étant en vertu de la formule (1) la coupe suivant $x_m = t$ de l'ensemble

$$h(g(C_0 \cap O') \cap g(\overline{V})) = (h \circ g)(C_0 \cap \overline{V}) = f(C_0 \cap \overline{V}),$$

il s'ensuit que toutes les coupes $f(C_0 \cap \overline{V})_t$ de l'ensemble $f(C_0 \cap \overline{V})$ sont des ensembles négligeables. D'après le théorème de Fubini, l'ensemble $f(C_0 \cap \overline{V})$ sera aussi négligeable. Ce qui prouve le lemme 1 pour $C_{s+1} = \varnothing$.

Supposons maintenant que $s > \frac{n}{m} - 1$. L'ensemble K étant recouvrable par un nombre fini de cubes de $\mathbb{R}^n$, puisque compact, il suffit de montrer que l'ensemble $f(C_s \cap Q)$ est négligeable pour tout cube $Q \subset \mathbb{R}^n$.

Soit a l'arête du cube Q et soit $k \geqslant 1$. Partageons le cube Q à l'aide de plans parallèles aux faces en k^n cubes d'arête a/k et considérons un cube Q' de cette partition, rencontrant C_s.

Les dérivées d'ordre $s + 1$ des fonctions $f^1, \ldots, f^m$ étant bornées dans Q, la formule de Taylor appliquée à ces fonctions nous dit que pour tous points $\mathbf{x} \in C_s \cap Q$ et $\mathbf{y} \in Q$

$$|f(\mathbf{y}) - f(\mathbf{x})| \leqslant M \cdot |\mathbf{y} - \mathbf{x}|^{s+1},$$

où M est une constante. Le diamètre du cube Q' étant visiblement égal à $\frac{a}{k}\sqrt{n}$, on en déduit que le diamètre de son image fQ' par f est $\leqslant M\left(\frac{a}{k}\sqrt{n}\right)^{s+1}$, donc son image est contenue dans un cube de $\mathbb{R}^m$ d'arête $2M\left(\frac{a}{k}\sqrt{n}\right)^{s+1}$ et de volume

$$\left[2M\left(\frac{a}{k}\sqrt{n}\right)^{s+1}\right]^m = \frac{B}{k^{(s+1)m}},$$

où $B = (2Ma^{s+1}(\sqrt{n})^{s+1})^m$ est une constante indépendante de k.

Puisque le nombre de tous les cubes Q' est $\leqslant k^n$ et que leur réunion contient l'ensemble $C_s \cap Q$, il s'ensuit que l'ensemble $f(C_s \cap Q)$ est contenu dans la réunion de cubes dont le volume total est inférieur à

$$\frac{B}{k^{(s+1)m}} k^n = \frac{B}{k^{(s+1)m-n}},$$

et donc tend vers 0 lorsque $k \to \infty$ (on rappelle que par hypothèse $(s + 1)m - n > 0$).

L'ensemble $f(C_s \cap Q)$ peut donc être recouvert par une famille (finie !) de cubes dont la somme des volumes est aussi petite que l'on veut. Cet ensemble est par conséquent de mesure nulle.

Ce qui prouve le lemme 1 et partant le théorème de Sard. □

Le théorème de Sard (ou plus exactement son corollaire) est la clef de la démonstration de la proposition 1 de la leçon précédente (ainsi que du théorème de Whitney), mais pour se servir de cette clef, il faut procéder à des constructions simples et intéressantes en soi.

* * *

Soient $\mathscr{X}$ et $\mathscr{Y}$ deux variétés (de dimension respective n et m) et soit $\mathscr{X} \times \mathscr{Y}$ l'ensemble de tous les couples (p, q), où $p \in \mathscr{X}$, $q \in \mathscr{Y}$. Pour tous ensembles $U \subset \mathscr{X}$ et $V \subset \mathscr{Y}$ l'ensemble $U \times V$ est une partie de l'ensemble $\mathscr{X} \times \mathscr{Y}$, et pour toutes applications $h : U \to \mathbb{R}^n$ et $k : V \to \mathbb{R}^m$ la formule

$$(h \times k)(p, q) = (h(p), k(q))$$

définit une application

$$h \times k : U \times V \to \mathbb{R}^{n+m}$$

(nous identifions $\mathbb{R}^n \times \mathbb{R}^m$ à $\mathbb{R}^{n+m}$). Ceci étant, si les applications h et k sont injectives, l'application $h \times k$ l'est aussi, et si les ensembles $h(U)$ et $k(V)$ sont ouverts (dans $\mathbb{R}^n$ et $\mathbb{R}^m$ respectivement), il en sera de même de l'ensemble

$$(h \times k)(U \times V) = h(U) \times k(V)$$

dans $\mathbb{R}^{n+m}$. Ceci exprime que si (U, h) et (V, k) sont des cartes, $(U \times V, h \times k)$ en sera une aussi. Bien plus, il est aisé de voir que *si* (U, h) *est compatible avec* (U', h'), *et* (V, k) *avec* (V', k'), *alors* $(U \times V, h \times k)$ *le sera avec* $(U' \times V', h' \times k')$. En effet, il est clair que

$$(U \times V) \cap (U' \times V') = (U \cap U') \times (V \cap V')$$

et de façon analogue

$$(h(U) \times k(V)) \cap (h'(U') \times k'(V')) = (h(U) \cap h'(U')) \times \\ \times (k(V) \cap k'(V'))$$

(on conviendra que $A \times B = \varnothing$ si $A = \varnothing$ ou $B = \varnothing$). Ceci étant,

$$(h \times k|_{(U \times V) \cap (U' \times V')}) \circ (h' \times k'|_{(U \times V) \cap (U' \times V')})^{-1} = \\ = [(h|_{U \cap U'}) \circ (h'|_{U \cap U'})^{-1}] \times [(k|_{V \cap V'}) \circ (k'|_{V \cap V'})^{-1}],$$

et, pour achever la démonstration, il reste à remarquer que pour tous difféomorphismes $\varphi : W \to W_1$ et $\varphi' : W' \to W'_1$ d'ensembles ouverts des espaces $\mathbb{R}^n$ et $\mathbb{R}^m$, l'application

$$\varphi \times \varphi' : W \times W' \to W_1 \times W'_1$$

est aussi un difféomorphisme d'ensembles ouverts (de l'espace $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$). □

Donc, les cartes $(U \times V,\ h \times k)$ construites pour toutes les cartes (U, h) et (V, k) des variétés $\mathcal{X}$ et $\mathcal{Y}$ forment un atlas sur $\mathcal{X} \times \mathcal{Y}$.

Définition 2. La structure différentiable sur $\mathcal{X} \times \mathcal{Y}$ s'appelle *produit direct des structures différentiables* des variétés $\mathcal{X}$ et $\mathcal{Y}$, et l'ensemble $\mathcal{X} \times \mathcal{Y}$, muni de cette structure, *produit direct des variétés $\mathcal{X}$ et $\mathcal{Y}$*. Sa dimension est égale à la somme des dimensions des facteurs :

$$\dim(\mathcal{X} \times \mathcal{Y}) = \dim \mathcal{X} + \dim \mathcal{Y}.$$

La topologie de la variété $\mathcal{X} \times \mathcal{Y}$ est visiblement le produit direct des topologies des variétés $\mathcal{X}$ et $\mathcal{Y}$ (cf. leçon 8).

Remarque 1. Pour tout groupe $\mathcal{G}$ est définie l'application

$$\mathcal{G} \times \mathcal{G} \to \mathcal{G}, \quad (a, b) \mapsto ab, \quad a, b \in \mathcal{G}. \tag{2}$$

Un groupe $\mathcal{G}$ qui est une variété différentiable pour laquelle l'application (2) est différentiable s'appelle *groupe de Lie* (ou encore *groupe différentiable*). La structure différentiable de tout groupe de Lie de matrices $\mathcal{G}$ (cf. définition 1 de la leçon 11) construite dans la proposition 1 de la leçon 11 possède la propriété suivante (vérifier-la !) : tout groupe $\mathcal{G}$ est un groupe de Lie pour cette structure différentiable. Ceci justifie la terminologie utilisée.

Remarque 2. Dans le prochain ouvrage on montrera que si un sous-groupe $\mathcal{G}$ du groupe $\mathrm{GL}(n, \mathbb{R})$ est une sous-variété plongée (donc, un groupe de Lie), il sera un groupe de Lie de matrices au sens de la définition 1 de la leçon 11. Ceci prouve que la définition 1 de la leçon 11 n'est pas, comme on pourrait le croire, une définition *ad hoc* et introduit une notion bien naturelle.

Si $\mathcal{X} = \mathcal{Y}$, on obtient une variété $\mathcal{X} \times \mathcal{X}$ appelée *carré* de $\mathcal{X}$. Le sous-ensemble Δ de cette variété, composé de tous les points de la forme (p, p), $p \in \mathcal{X}$, s'appelle *diagonale* de $\mathcal{X} \times \mathcal{X}$.

Il est aisé de voir qu'*une variété $\mathcal{X}$ (ou d'une façon plus générale un espace topologique $\mathcal{X}$) est séparée si et seulement si la diagonale Δ est fermée dans $\mathcal{X} \times \mathcal{X}$*. En effet, dire que des points distincts p et q ont des voisinages U et V disjoints revient exactement à dire que le voisinage $U \times V$ du point $(p, q) \in \mathcal{X} \times \mathcal{X}$ ne rencontre pas Δ. $\square$

Donc, pour toute variété séparée $\mathcal{X}$ est définie la variété $\mathcal{X} \times \mathcal{X} \setminus \Delta$ de dimension $2n$, où $n = \dim \mathcal{X}$.

* * *

L'autre construction associe à une variété $\mathcal{X}$ de dimension n une variété $\mathrm{T}\mathcal{X}$ de dimension $2n$.

Soit

$$\mathrm{T}\mathcal{X} = \bigsqcup_{p \in \mathcal{X}} \mathrm{T}_p \mathcal{X}$$

la réunion disjointe de tous les sous-espaces $\mathrm{T}_p\mathscr{X}$, $p \in \mathscr{X}$. [Les points de l'ensemble $\mathrm{T}\mathscr{X}$ seront ainsi les vecteurs A tangents à $\mathscr{X}$.] Pour tout vecteur $A \in \mathrm{T}\mathscr{X}$ on désignera par πA le point $p \in \mathscr{X}$ (unique !) tel que $A \in \mathrm{T}_p\mathscr{X}$. Il apparaît ainsi une application

$$\pi : \mathrm{T}\mathscr{X} \to \mathscr{X}$$

telle que $\pi^{-1}(p) = \mathrm{T}_p\mathscr{X}$ pour tout point $p \in \mathscr{X}$.

Pour tout sous-ensemble ouvert $U \subset \mathscr{X}$, le sous-ensemble $\pi^{-1}U \subset \subset \mathrm{T}\mathscr{X}$ s'identifie de façon naturelle à l'ensemble $\mathrm{T}U$. Lorsque U est le support d'une carte $(U, h) = (U, x^1, \ldots, x^n)$, est définie l'application

$$\mathrm{T}h : \mathrm{T}U \to \mathbb{R}^{2n},$$

associant à un vecteur arbitraire $A \in \mathrm{T}U$ le vecteur

$$(\mathrm{T}h)\,A = (x^1, \ldots, x^n, a^1, \ldots, a^n) \in \mathbb{R}^{2n},$$

où $x^1, \ldots, x^n$ sont les coordonnées du point $p = \pi A$ dans la carte (U, h), et $a^1, \ldots, a^n$ les coordonnées du vecteur A dans la base

$$\left(\frac{\partial}{\partial x^1}\right)_p, \ \ldots, \ \left(\frac{\partial}{\partial x^n}\right)_p$$

de l'espace $\mathrm{T}_p\mathscr{X}$. (Donc

$$A = a^1\left(\frac{\partial}{\partial x^1}\right)_p + \cdots + a^n\left(\frac{\partial}{\partial x^n}\right)_p;$$

on remarquera que les nombres $a^1, \ldots, a^n$ sont définis de façon unique par le vecteur A.) Il est clair que l'application $\mathrm{T}h$ est bijective et donc le couple $(\mathrm{T}U, \mathrm{T}h)$ est une carte dans $\mathrm{T}\mathscr{X}$.

Soient (U, h) et (U', h') deux cartes de la variété $\mathscr{X}$ (pour fixer les idées, des cartes sécantes) et soit

$$x^{i'} = x^{i'}(x^1, \ldots, x^n), \quad i = 1, \ldots, n, \tag{3}$$

les formules de passage des coordonnées locales correspondantes (dans l'intersection $U \cap U'$). Par définition (cf. formule (1) de la leçon 12) pour tout point $p \in U \cap U'$ les coordonnées $a^1, \ldots, a^n$ et $a^{1'}, \ldots, a^{n'}$ de tout vecteur $A \in \mathrm{T}_p\mathscr{X}$ dans les cartes (U, h) et (U', h') sont reliées par

$$a^{i'} = \left(\frac{\partial x^{i'}}{\partial x^i}\right)_p a^i, \tag{4}$$

où $\left(\frac{\partial x^{i'}}{\partial x^i}\right)_p$ sont les valeurs des dérivées partielles des fonctions (3) au point p.

D'autre part il est clair que $\mathrm{T}h$ et $\mathrm{T}h'$ appliquent l'ensemble

$$\mathrm{T}(U \cap U') = \mathrm{T}U \cap \mathrm{T}U'$$

respectivement sur les ensembles ouverts $h(U \cap U') \times \mathbb{R}^n$ et $h'(U \cap U') \times \mathbb{R}^n$ de l'espace $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$, et de plus les

formules (3) et (4) définissent l'application $\mathrm{T}h' \circ (\mathrm{T}h)^{-1}$ du premier ensemble sur le second. Cette application est un difféomorphisme puisque le jacobien $\det \left\| \left(\frac{\partial x^{i'}}{\partial x^i} \right)_p \right\|$ est différent de 0 en tout point $p \in U \cap U'$.

Donc, *les cartes* $(\mathrm{T}U, \mathrm{T}h)$ *et* $(\mathrm{T}U', \mathrm{T}h')$ *sont compatibles*. Ceci vaut également dans le cas où $U \cap U' = \varnothing$ (car si $U \cap U' = \varnothing$, alors $\mathrm{T}U \cap \mathrm{T}U' = \varnothing$).

Donc, les cartes $(\mathrm{T}U, \mathrm{T}h)$ forment un atlas sur $\mathrm{T}\mathscr{X}$ et, par suite, définissent une structure différentiable sur $\mathrm{T}\mathscr{X}$.

Définition 3. La variété différentiable $\mathrm{T}\mathscr{X}$ s'appelle *variété des vecteurs tangents* à la variété $\mathscr{X}$. Elle est de dimension $2n$, où $n = \dim \mathscr{X}$.

A noter que la présence des dérivées dans la formule (4) explique que la classe de différentiabilité de la variété $\mathrm{T}\mathscr{X}$ est d'une unité inférieure à celle de la variété $\mathscr{X}$ (si $r = \infty$ ou $r = \omega$, les deux variétés sont visiblement de même classe C^r).

Par définition, les coordonnées locales correspondant à la carte $(\mathrm{T}U, \mathrm{T}h)$ sont les nombres $x^1, \ldots, x^n, a^1, \ldots, a^n$. (Donc, les symboles $x^1, \ldots, x^n$ désignent des coordonnées locales simultanément dans U et dans $\mathrm{T}U$. Un peu d'attention suffit pour éviter toute ambiguïté.)

Dans les coordonnées locales $x^1, \ldots, x^n, a^1, \ldots, a^n$ (sur $\mathrm{T}U$) et $x^1, \ldots, x^n$ (sur U) l'application π est définie par les formules

$$x^i = x^i, \quad i = 1, \ldots, n,$$

où, au premier membre, x^i sont des coordonnées sur U, et, au second, sur $\mathrm{T}U$. On voit donc que *l'application* π *est différentiable et est une submersion*. Par conséquent, en vertu de la proposition générale 1 de la leçon 13, *chaque espace tangent* $\mathrm{T}_p\mathscr{X} = \pi^{-1}(p)$ *est une sous-variété plongée dans* $\mathrm{T}\mathscr{X}$. Les nombres $a^1, \ldots, a^n$ sont des coordonnées sur cette sous-variété (définies sur $\mathrm{T}_p\mathscr{X}$).

La variété $\mathrm{T}\mathscr{X}$ contient un sous-ensemble $\mathscr{X}_0$ composé des vecteurs nuls des espaces $\mathrm{T}_p\mathscr{X}$.

Il est aisé de voir que $\mathscr{X}_0$ *est une sous-variété fermée difféomorphe à* $\mathscr{X}$ (le difféomorphisme $\mathscr{X}_0 \to \mathscr{X}$ est induit par l'application $\pi : \mathrm{T}\mathscr{X} \to \mathscr{X}$).

L'ensemble $\mathrm{T}\mathscr{X} \setminus \mathscr{X}_0$ est ouvert, donc c'est une variété dont la dimension est $2n$, où $n = \dim \mathscr{X}$.

* * *

Nous pouvons désormais prouver la proposition 1 de la leçon 14 (et avec elle le théorème de Whitney).

D é m o n s t r a t i o n d e l a p r o p o s i t i o n 1 d e l a l e ç o n 14. Supposons que la variété $\mathscr{X}$ est plongée dans $\mathbb{R}^N$. Pour

tout couple de points distincts $p, q \in \mathscr{X}$ (i.e. pour tout point $(p, q) \in \in \mathscr{X} \times \mathscr{X} \setminus \Delta$) est définie une droite de $\mathbb{R}^N$ passant par ces points. Soit $f_1(p, q)$ le sous-espace à une dimension de $\mathbb{R}^N$ associé à cette droite (i.e. de façon suggestive, la droite parallèle passant par 0). Puisque les sous-espaces à une dimension de $\mathbb{R}^N$ engendrent l'espace projectif $\mathbb{R}P^{N-1}$, ceci nous donne l'application

$$f_1 : \mathscr{X} \times \mathscr{X} \setminus \Delta \to \mathbb{R}P^{N-1}. \tag{5}$$

On définit de façon analogue l'application

$$f_2 : \mathrm{T}\mathscr{X} \setminus \mathscr{X}_0 \to \mathbb{R}P^{N-1}, \tag{6}$$

en prenant pour $f_2(A)$, $\forall A \in \mathrm{T}\mathscr{X} \setminus \mathscr{X}_0$, le sous-espace à une dimension de $\mathbb{R}^N$ (i.e. le point de l'espace $\mathbb{R}P^{N-1}$) engendré par A (ou plus exactement par le vecteur $(d\iota_p)A \in \mathrm{T}_p\mathbb{R}^N = \mathbb{R}^N$, où $d\iota_p$ est la différentielle de l'injection $\iota : \mathscr{X} \to \mathbb{R}^N$ au point $p = \pi(A)$).

Exercice 1. Montrer que les applications (5) et (6) sont différentiables. [N o t a. Représenter ces applications en coordonnées locales.]

Soit maintenant

$$\mathscr{Y} = (\mathscr{X} \times \mathscr{X} \setminus \Delta) \sqcup (\mathrm{T}\mathscr{X} \setminus \mathscr{X}_0)$$

la réunion disjointe des variétés $\mathscr{X} \times \mathscr{X} \setminus \Delta$ et $\mathrm{T}\mathscr{X} \setminus \mathscr{X}_0$. Cette réunion est une variété différentiable à $2n$ dimensions, et l'application

$$f : \mathscr{Y} \to \mathbb{R}P^{N-1},$$

qui est confondue avec f_1 sur $\mathscr{X} \times \mathscr{X} \setminus \Delta$ et avec f_2 sur $\mathrm{T}\mathscr{X} \setminus \mathscr{X}_0$, est une application différentiable. Donc, si $2n < N - 1$, l'application f n'est visiblement pas surjective en vertu du corollaire au théorème de Sard. Ceci prouve que *dans l'espace* $\mathbb{R}^N$ *il existe un sous-espace* L *à une dimension n'appartenant ni à* $\mathrm{Im}\, f_1$ *ni à* $\mathrm{Im}\, f_2$.

Considérons maintenant la projection

$$\mathbb{R}^N \to L^\perp \tag{7}$$

de l'espace $\mathbb{R}^N$ parallèlement à la droite L sur son orthogonal $L^\perp$ (qui est un sous-espace à $N - 1$ dimensions de $\mathbb{R}^N$). En choisissant une base dans $L^\perp$, on peut admettre que cette projection est une application de $\mathbb{R}^N$ dans $\mathbb{R}^{N-1}$. Soit

$$\rho : \mathscr{X} \to \mathbb{R}^{N-1} \tag{8}$$

la restriction de cette application à $\mathscr{X} \subset \mathbb{R}^N$. La projection (7) étant continue et ouverte, sa restriction (8) l'est aussi. D'autre part, dire que L n'appartient pas à $\mathrm{Im}\, f_1$ revient à dire que toute droite parallèle à L coupe la sous-variété $\mathscr{X}$ en un point au plus, i.e. que

l'application (8) est injective. Etant une application injective continue et ouverte, l'application (8) est un monéomorphisme.

Considérons maintenant la condition $L \notin \operatorname{Im} f_2$ (qui géométriquement exprime soit dit en passant que la droite parallèle à L et qui passe par p n'est pas tangente à $\mathscr{X}$ en ce point). En tout point p de $\mathscr{X}$ la différentielle $(d\rho)_p$ de l'application (8) est visiblement la restriction à $T_p\mathscr{X} \subset \mathbb{R}^N$ de la différentielle de la projection (7). L'application (7) étant linéaire, sa différentielle est confondue avec elle, d'où il s'ensuit en particulier que la différentielle $(d\rho)_p$ n'annule que les vecteurs de $T_p\mathscr{X} \cap L$. Or cette intersection soit est constituée du seul vecteur nul (lorsque $L \not\subset T_p\mathscr{X}$), soit est confondue avec L (lorsque $L \subset T_p\mathscr{X}$). Donc, si $L \not\subset T_p\mathscr{X}$, l'application (8) est une immersion au point p. Puisque l'inclusion $L \subset T_p\mathscr{X}$ exprime que $L = f_2(p)$, ceci prouve que la condition $L \notin \operatorname{Im} f_2$ équivaut au fait que la projection est une immersion (en chaque point $p \in \mathscr{X}$) et donc est un difféomorphisme sur son image.

Donc, l'image $\rho(\mathscr{X})$ de la variété $\mathscr{X}$ par la projection ρ est une sous-variété plongée dans l'espace $\mathbb{R}^{N-1}$, difféomorphe à la variété $\mathscr{X}$. Donc, la variété $\mathscr{X}$ est plongeable dans $\mathbb{R}^{N-1}$. □

Exercice 2. Montrer que toute variété à n dimensions plongeable dans $\mathbb{R}^N$ peut être immergée dans $\mathbb{R}^{2n}$.

LEÇON 16

Tenseurs. — Champs de tenseurs. — Champs de vecteurs et dérivation. — Algèbre de Lie de champs de vecteurs.

On rappelle (cf. leçon II.5) qu'un *tenseur S de type* (a, b), où $a \geqslant 0$, $b \geqslant 0$, sur un espace vectoriel $\mathscr{V}$ est par définition une application associant à une base arbitraire $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de $\mathscr{V}$ une collection de n^{a+b} nombres $S^{j_1 \ldots j_b}_{i_1 \ldots i_a}$ appelés *composantes* de S dans cette base et tels que les composantes de S dans des bases quelconques $\mathbf{e}_1, \ldots, \mathbf{e}_n$ et $\mathbf{e}_{1'}, \ldots, \mathbf{e}_{n'}$ de $\mathscr{V}$ soient reliées par la formule

$$S^{j'_1 \ldots j'_b}_{i'_1 \ldots i'_a} = c^{i_1}_{i'_1} \ldots c^{i_a}_{i'_a} c^{j'_1}_{j_1} \ldots c^{j'_b}_{j_b} S^{j_1 \ldots j_b}_{i_1 \ldots i_a},$$

où $c^{i'}_i$ et $c^i_{i'}$ sont les éléments de matrices de passage inverses l'une de l'autre, i.e. des nombres tels que $\mathbf{e}_{i'} = c^i_{i'} \mathbf{e}_i$ et $\mathbf{e}_i = c^{i'}_i \mathbf{e}_{i'}$. Chaque tenseur définit de façon unique une fonctionnelle multilinéaire

$$S(\mathbf{x}_1, \ldots, \mathbf{x}_a, \xi^1, \ldots, \xi^b) = S^{j_1 \ldots j_b}_{i_1 \ldots i_a} x^{i_1}_1 \ldots x^{i_a}_a \xi^1_{j_1} \ldots \xi^b_{j_b}$$

de a vecteurs et b covecteurs et, d'une façon générale, s'identifie à cette fonctionnelle.

Les tenseurs de type (a, b) forment un espace vectoriel $\mathrm{T}^b_a(\mathscr{V})$ par rapport aux opérations d'addition et de multiplication par un nombre, définies de façon naturelle.

Pour tous tenseurs S et R de types respectifs (a, b) et (c, d) la formule

$$(S \otimes R)^{j_1 \ldots j_{b+d}}_{i_1 \ldots i_{a+c}} = S^{j_1 \ldots j_b}_{i_1 \ldots i_a} R^{j_{b+1} \ldots j_{b+d}}_{i_{a+1} \ldots i_{a+c}}$$

définit un tenseur $S \otimes R$ de type $(a + c, b + d)$ appelé *produit tensoriel* des tenseurs S et R. Cette multiplication est associative et distributive par rapport à l'addition.

De plus, les tenseurs sont justiciables d'une opération spéciale appelée *convolution* (cf. leçon II.6).

Tout vecteur est assimilé à un tenseur de type (0, 1), et tout covecteur, à un tenseur de type (1, 0). Donc, pour toute base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de $\mathcal{V}$ et tous indices $i_1, \ldots, i_a, j_1, \ldots, j_b = 1, \ldots, n$, est défini dans l'espace $\mathrm{T}_a^b(\mathcal{V})$ le tenseur

$$\mathbf{e}^{i_1} \otimes \ldots \otimes \mathbf{e}^{i_a} \times \mathbf{e}_{j_1} \otimes \ldots \otimes \mathbf{e}_{j_b},$$

où $\mathbf{e}^{i_1}, \ldots, \mathbf{e}^{i_a}$ sont des vecteurs de la base duale de l'espace $\mathcal{V}^*$.

Les tenseurs de cette forme engendrent une base de l'espace $\mathrm{T}_a^b(\mathcal{V})$ dans laquelle les coordonnées de tout tenseur S sont confondues avec ses composantes, i.e.

$$S = S_{i_1 \ldots i_a}^{j_1 \ldots j_b} \mathbf{e}^{i_1} \otimes \ldots \otimes \mathbf{e}^{i_a} \otimes \mathbf{e}_{j_1} \otimes \ldots \otimes \mathbf{e}_{j_b}.$$

* * *

Nous appliquerons ces notions générales d'algèbre linéaire au cas où $\mathcal{V}$ est un espace tangent $\mathrm{T}_p\mathcal{X}$ à une variété différentiable $\mathcal{X}$ en un point p.

Soit $(U, h) = (U, x^1, \ldots, x^n)$ une carte arbitraire d'une variété $\mathcal{X}$ et soit p un point de cette carte. Dans l'espace $\mathrm{T}_p\mathcal{X}$ cette carte définit la base

$$\left(\frac{\partial}{\partial x^1}\right)_p, \ \ldots, \ \left(\frac{\partial}{\partial x^n}\right)_p, \tag{1}$$

et dans l'espace dual $\mathrm{T}_p^*\mathcal{X}$, la base duale

$$(\mathrm{d}x^1)_p, \ \ldots, \ (\mathrm{d}x^n)_p.$$

Donc, sur l'espace $\mathrm{T}_p\mathcal{X}$, tout tenseur S_p de type (a, b) se représente sous la forme

$$\begin{aligned} S_p &= \\ &= S_{i_1 \ldots i_a}^{j_1 \ldots j_b} (\mathrm{d}x^{i_1})_p \otimes \ldots \otimes (\mathrm{d}x^{i_a})_p \otimes \left(\frac{\partial}{\partial x^{j_1}}\right)_p \otimes \ldots \otimes \left(\frac{\partial}{\partial x^{j_b}}\right)_p. \end{aligned} \tag{2}$$

Les coefficients $S_{i_1 \ldots i_a}^{j_1 \ldots j_b}$ (i.e. les composantes du tenseur S_p dans la base (1)) s'appellent *composantes du tenseur S_p dans la carte (U, h)*. (Pour des raisons d'ordre typographique nous omettrons l'indice p dans la notation de ces composantes.)

Toute autre carte $(U', h') = (U', x^{1'}, \ldots, x^{n'})$ (avec $p \in U'$) définit une base

$$\left(\frac{\partial}{\partial x^{1'}}\right)_p, \ \ldots, \ \left(\frac{\partial}{\partial x^{n'}}\right)_p \tag{3}$$

dans l'espace $\mathrm{T}_p\mathcal{X}$, liée à la base (1) par la matrice de passage

$$\left\| \left(\frac{\partial x^{i'}}{\partial x^i}\right)_p \right\|, \quad i, i' = 1, \ldots, n. \tag{4}$$

Donc, les composantes du tenseur S_p dans les cartes (U, h) et (U', h') sont reliées par

$$(5) \qquad S^{j'_1 \cdots j'_b}_{i'_1 \cdots i'_a} = \left(\frac{\partial x^{i_1}}{\partial x^{i'_1}}\right)_p \cdots \left(\frac{\partial x^{i_a}}{\partial x^{i'_a}}\right)_p \left(\frac{\partial x^{j'_1}}{\partial x^{j_1}}\right)_p \cdots \left(\frac{\partial x^{j'_b}}{\partial x^{j_b}}\right)_p S^{j_1 \cdots j_b}_{i_1 \cdots i_a}.$$

Si maintenant le tenseur S_p est défini pour tout point $p \in \mathcal{X}$, dans la représentation (2) les composantes $S^{j_1 \cdots j_b}_{i_1 \cdots i_a}$ seront des fonctions de p. Si ces fonctions sont différentiables, i.e. s'expriment différentiablement dans la carte (U, h) au moyen des coordonnées $x^1, \ldots, x^n$, la correspondance $p \mapsto S_p$ s'appelle *champ (différentiable) de tenseurs* (ou plus brièvement *tenseur*) *de type* (a, b) *sur la variété* $\mathcal{X}$. Pour les fonctions $S^{j_1 \cdots j_b}_{i_1 \cdots i_a}$ la relation (5) devient sur $U \cap U'$

$$(6) \qquad S^{j'_1 \cdots j'_b}_{i'_1 \cdots i'_a} = \frac{\partial x^{i_1}}{\partial x^{i'_1}} \cdots \frac{\partial x^{i_a}}{\partial x^{i'_a}} \frac{\partial x^{j'_1}}{\partial x^{j_1}} \cdots \frac{\partial x^{j'_b}}{\partial x^{j_b}} S^{j_1 \cdots j_b}_{i_1 \cdots i_a},$$

d'où il résulte que la condition de différentiabilité d'un champ de tenseurs est indépendante du choix de la carte.

Remarque 1. Si les variétés envisagées sont de classe C^r, où r est un nombre fini $\geqslant 1$, on se heurte à une difficulté caractéristique due au fait que les éléments de la matrice (4) ne sont en général que de classe C^{r-1}. Ceci nous contraint à étudier des champs de tenseurs de classe C^{r-1}. C'est pour éviter ces restrictions qu'on a convenu à la leçon 6 de se limiter aux variétés de classe C^∞ (ou C^ω) pour lesquelles cette difficulté est exclue.

Pour tout recouvrement ouvert $\{U_\alpha\}$ de la variété $\mathcal{X}$, chaque champ de tenseurs S définit une famille de champs

$$S_\alpha = S|_{U_\alpha}$$

tels que pour tous indices α et β

$$(7) \qquad S_\alpha = S_\beta \quad \text{sur} \quad U_\alpha \cap U_\beta.$$

Réciproquement, si sont donnés des champs S_α sur U_α satisfaisant aux relations (7) (on dira de ces champs qu'ils sont *compatibles sur les intersections*), la formule

$$S_p = (S_\alpha)_p, \quad p \in U_\alpha,$$

définit de façon unique sur $\mathcal{X}$ un champ de tenseurs S tel que

$$S|_{U_\alpha} = S_\alpha$$

pour tout α (un champ qui est donc différentiable). On dira que les champs S_α *composent* S.

Remarque 2. Un champ de tenseurs sur une variété $\mathscr{X}$ peut être traité comme une correspondance associant à chaque carte (U, h) de $\mathscr{X}$ une collection de fonctions différentiables $S^{j_1 \cdots j_b}_{i_1 \cdots i_a}$ sur U telles que la relation (6) ait lieu sur l'intersection $U \cap U'$ des supports de deux cartes quelconques (U, h) et (U', h'). Ceci peut être adopté pour d é f i n i t i o n d'un champ de tenseurs. Cette définition présente l'avantage de pouvoir être formulée immédiatement après l'introduction de la notion de variété différentiable sans définitions intermédiaires, et l'inconvénient de ne pas avoir un lien formel direct (substituable à une analogie) avec la notion de tenseur dans un espace vectoriel.

Toutes les opérations algébriques sur les tenseurs (y compris la convolution) se généralisent automatiquement aux champs de tenseurs. Ainsi, le produit tensoriel $S \otimes R$ de deux champs de tenseurs S et R est défini par la formule

$$(S \otimes R)_p = S_p \otimes R_p. \tag{8}$$

Il est clair que des champs différentiables donneront des champs différentiables par ces opérations.

On voit en particulier que *l'ensemble* $\mathsf{T}^b_a\mathscr{X}$ *des champs de tenseurs de type* (a, b) *sur une variété* $\mathscr{X}$ *est muni d'une structure d'espace vectoriel.*

Cet espace est de dimension infinie (pour $n > 0$).

Pour $(a, b) = (0, 0)$ les champs de tenseurs ne sont autres que des fonctions différentiables sur $\mathscr{X}$, et l'espace vectoriel $\mathsf{T}^0_0\mathscr{X}$, l'espace vectoriel $\mathsf{F}\mathscr{X}$ des fonctions différentiables sur $\mathscr{X}$. L'espace vectoriel $\mathsf{F}\mathscr{X}$ est une algèbre par rapport à la multiplication des fonctions, et de plus la formule

$$(fS)_p = f(p)\, S_p, \quad f \in \mathsf{F}\mathscr{X}, \quad S \in \mathsf{T}^b_a\mathscr{X}$$

(qui est un cas particulier de la formule (8)) définit l'opération de multiplication

$$\mathsf{F}\mathscr{X} \times \mathsf{T}^b_a\mathscr{X} \to \mathsf{T}^b_a\mathscr{X}$$

par rapport à laquelle, ainsi que le montre une vérification immédiate, *l'espace vectoriel* $\mathsf{T}^b_a\mathscr{X}$ *est un module sur l'algèbre* $\mathsf{F}\mathscr{X}$.

Pour $(a, b) = (0, 1)$ les champs de tenseurs s'appellent *champs de vecteurs.* Un exemple de champs de vecteurs sur un voisinage de coordonnées U (traité comme une variété) nous est fourni par le champ :

$$\frac{\partial}{\partial x^i} : p \mapsto \left(\frac{\partial}{\partial x^i}\right)_p, \quad i = 1, \ldots, n, \tag{9}$$

qui s'appelle *i-ième champ de vecteurs de coordonnées* sur U.

Pour $(a, b) = (1, 0)$ les champs de tenseurs s'appellent *champs de covecteurs*. Exemple: le *i-ième champ de covecteurs de coordonnées*

$$\mathrm{d}x^i : p \mapsto (\mathrm{d}x^i)_p \tag{10}$$

sur le voisinage de coordonnées U.

La formule (2) affirme que tout champ de tenseurs S sur U se décompose d'une seule façon en un produit tensoriel de champs de vecteurs et de covecteurs de coordonnées sur U:

$$S = S^{j_1 \ldots j_b}_{i_1 \ldots i_a} \mathrm{d}x^{i_1} \otimes \ldots \otimes \mathrm{d}x^{i_a} \otimes \frac{\partial}{\partial x^{j_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{j_b}}. \tag{11}$$

En particulier, sur U tout champ de vecteurs X est de la forme

$$X = X^i \frac{\partial}{\partial x^i}. \tag{12}$$

et tout champ de covecteurs α, de la forme

$$\alpha = \alpha_i \, \mathrm{d}x^i, \tag{13}$$

où X^i et α_i, $i = 1, \ldots, n$, sont des fonctions différentiables sur U.

Par définition, l'existence du développement (11) exprime que *pour tout voisinage de coordonnées U, l'espace vectoriel $\mathsf{T}^b_a U$ est un module libre sur l'algèbre $\mathsf{F}U$ de base*

$$\mathrm{d}x^{i_1} \otimes \ldots \otimes \mathrm{d}x^{i_a} \otimes \frac{\partial}{\partial x^{j_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{j_b}}.$$

Si $\mathscr{X}$ est une variété arbitraire, le module $\mathsf{T}^b_a \mathscr{X}$ n'est généralement pas libre (sur l'algèbre $\mathsf{F}\mathscr{X}$) et sa structure algébrique peut être très compliquée.

Une variété $\mathscr{X}$ dont tous les modules $\mathsf{T}^b_a \mathscr{X}$ sont libres s'appelle *parallélisable*.

* * *

Etudions les champs de vecteurs plus en détail.

Comme déjà dit à la leçon 12, tout vecteur $A \in \mathrm{T}_p \mathscr{X}$ permet d'associer à une fonction f (définie et différentiable au voisinage de p) un nombre Af, dérivée de cette fonction par rapport à A. Il s'ensuit que pour tout champ de vecteurs X sur une variété $\mathscr{X}$ et toute fonction $f \in \mathsf{F}\mathscr{X}$, la formule

$$(Xf)(p) = X_p f, \quad p \in \mathscr{X}, \tag{14}$$

définit une fonction Xf sur $\mathscr{X}$. Les formules pour Af citées à la leçon 12 impliquent que dans une carte arbitraire $(U, h) = (U, x^1, \ldots, x^n)$ de $\mathscr{X}$ la restriction de la fonction Xf à U est

définie par la formule

$$Xf = X^i \frac{\partial f}{\partial x^i} \quad \text{sur } U, \tag{15}$$

où X^i, $i = 1, \ldots, n$, sont les composantes du champ de vecteurs X dans la carte (U, h). Donc, la fonction Xf est différentiable sur U et, par suite, sur $\mathscr{X}$ tout entière, puisque U est arbitraire.

La formule (14) définit par conséquent une application X (visiblement linéaire) de $\mathsf{F}\mathscr{X}$ dans $\mathsf{F}\mathscr{X}$, appelée *opérateur différentiel linéaire du premier ordre* sur $\mathscr{X}$, engendré par le champ de vecteurs X. [Cette terminologie est motivée par la formule (15) qui, comparée à la formule (12), explique aussi le choix de la notation $\frac{\partial}{\partial x^i}$ pour les champs de vecteurs de coordonnées.]

Soit $\mathscr{A}$ une algèbre (pas nécessairement de dimension finie et associative).

Définition 1. Une application linéaire

$$D : \mathscr{A} \to \mathscr{A}$$

s'appelle *dérivation* si

$$D(ab) = Da \cdot b + a \cdot Db$$

pour tous éléments $a, b \in \mathscr{A}$.

En particulier, les dérivations de l'algèbre $\mathsf{F}\mathscr{X}$ (appelées généralement tout simplement *dérivations sur* $\mathscr{X}$) sont des applications linéaires

$$D : \mathsf{F}\mathscr{X} \to \mathsf{F}\mathscr{X}$$

telles que

$$D(fg) = Df \cdot g + f \cdot Dg \tag{16}$$

pour des fonctions f et g quelconques différentiables sur $\mathscr{X}$.

Il est aisé de voir que *l'opérateur différentiel linéaire* $X : \mathsf{F}\mathscr{X} \to \mathsf{F}\mathscr{X}$ *engendré par un champ de vecteurs* X *est une dérivation sur* $\mathscr{X}$. En effet, la règle de dérivation du produit et la formule (15) entraînent la réalisation immédiate de l'identité (16) sur chaque voisinage de coordonnées U pour toutes fonctions $f, g \in \mathsf{F}\mathscr{X}$. Donc, elle est réalisée sur la variété $\mathscr{X}$ tout entière. □

Il se trouve que ceci épuise toutes les dérivations sur $\mathscr{X}$ si la variété $\mathscr{X}$ est séparée.

Théorème 1. *Toute dérivation D sur une variété $\mathscr{X}$ séparée différentiable (de classe C^∞) est engendrée par un champ de vecteurs. Ce champ est unique.*

Pour prouver ce théorème nous aurons besoin du lemme suivant:

Lemme 1. *Soient $\mathscr{X}$ une variété séparée différentiable, U une sous-variété ouverte de $\mathscr{X}$, f une fonction différentiable sur U. Pour tout*

point $p_0 \in U$ il existe alors sur $\mathscr{X}$ une fonction différentiable f_1 et un voisinage W de p_0 tels que $\overline{W} \subset U$ et

$$f = f_1 \quad \textit{sur } W.$$

Ceci étant, on peut admettre subsidiairement que $f_1 = 0$ à l'extérieur de U.

Démonstration. En vertu de la proposition 2 de la leçon 14, il existe dans $\mathscr{X}$ des ensembles ouverts V et W tels que

$$p_0 \in W, \quad \overline{W} \subset V, \quad \overline{V} \subset U,$$

et pour le couple (V, W) existe la fonction d'Urysohn φ. Pour tout point $p \in \mathscr{X}$ posons

$$f_1(p) = \begin{cases} \varphi(p) f(p) & \text{si } p \in U, \\ 0 & \text{si } p \notin U. \end{cases}$$

Il est clair que la fonction f_1 est différentiable et coïncide avec la fonction f sur W. D'autre part, par construction, $f_1 = 0$ à l'extérieur de U. □

Corollaire 1. *Pour tout voisinage U d'un point arbitraire p_0 d'une variété séparée différentiable $\mathscr{X}$, il existe un voisinage W de p_0 et une fonction φ différentiable sur $\mathscr{X}$ tels que*

$$\varphi(p) = \begin{cases} 1 & \textit{si } p \in W, \\ 0 & \textit{si } p \notin U. \end{cases} \quad \square$$

On remarquera que $\varphi(p_0) = 1$ et $\overline{W} \subset U$.

Soit maintenant D une dérivation arbitraire sur une variété $\mathscr{X}$.

On dira que des fonctions f et g différentiables sur $\mathscr{X}$ sont *égales à proximité d'un point $p_0 \in \mathscr{X}$* si elles prennent les mêmes valeurs dans un voisinage de p_0.

Corollaire 2. *Si des fonctions f et g sont égales à proximité d'un point de $\mathscr{X}$, il en sera de même des fonctions Df et Dg.*

Démonstration. Supposons que $f = g$ dans un voisinage U du point p_0. Considérons la fonction $(f - g)\varphi$, où φ est la fonction du corollaire 1. Il est clair que $(f - g)\varphi = 0$ sur $\mathscr{X}$ tout entière, i.e. la fonction $(f - g)\varphi$ est le zéro de l'espace vectoriel $\mathsf{F}\mathscr{X}$. Donc, la fonction $D[(f - g)\varphi]$ est aussi un zéro, puisque D est linéaire. Comme

$$D[(f - g)\varphi] = (Df - Dg)\varphi + (f - g)D\varphi,$$

ceci prouve que

$$(Df - Dg)\varphi + (f - g)D\varphi = 0$$

sur $\mathscr{X}$ tout entière et en particulier sur le voisinage W de p_0 envisagé dans le corollaire 1. Or, $f = g$ et $\varphi = 1$ sur W, donc $Df - Dg = 0$ et, par suite, $Df = Dg$ sur W. □

Remarque 3. Du lemme 1 il s'ensuit également que *pour tout point $p_0 \in U$ et tout champ de tenseurs S sur U il existe un voisinage $W \subset U$ de p_0 et un champ de tenseurs S_1 sur $\mathscr{X}$ tout entière, tels que*

$$S = S_1 \quad sur\ W.$$

Pour prouver ceci, il suffit d'appliquer le lemme 1 à chaque composante du champ S (et à prendre l'intersection des voisinages correspondants W).

La propriété exprimée par le corollaire 2 s'appelle p r o p r i é t é l o c a l e de l'application D. Cette propriété entraîne l'importante proposition suivante:

Proposition 1. *Pour tout ensemble ouvert $U \subset \mathscr{X}$ il existe une seule dérivation*

$$D_U : \mathsf{F}U \to \mathsf{F}U$$

compatible avec D, i.e. telle que

$$D_U (f\,|_U) = Df\,|_U \tag{17}$$

pour toute fonction f différentiable sur $\mathscr{X}$.

De façon suggestive cette propriété exprime que dans le diagramme

$$\begin{array}{ccc} \mathsf{F}\mathscr{X} & \to & \mathsf{F}U \\ {\scriptstyle D}\downarrow & & \downarrow{\scriptstyle D_U}, \\ \mathsf{F}\mathscr{X} & \to & \mathsf{F}U \end{array}$$

dont les flèches horizontales figurent des restrictions, on obtient la même application $\mathsf{F}\mathscr{X} \to \mathsf{F}U$ quel que soit le chemin suivi pour aller du coin supérieur gauche au coin inférieur droit. De tels diagrammes sont dits *commutatifs* (nous avons eu affaire aux diagrammes commutatifs à la leçon 3).

D é m o n s t r a t i o n d e l a p r o p o s i t i o n 1. Supposons que la dérivation D_U existe et soit g une fonction différentiable sur U. En vertu du lemme 1, pour tout point $p_0 \in U$ il existe sur $\mathscr{X}$ une fonction $g_1 = g_{1,p_0}$ confondue avec g à proximité de p_0. En outre, en vertu de la propriété (17), $D_U g = Dg_1\,|_U$ et donc en particulier

$$(D_U g)\,(p_0) = (Dg_1)\,(p_0). \tag{18}$$

D'autre part, si g_1^* est une fonction différentiable sur $\mathscr{X}$, également confondue avec g (donc avec g_1) à proximité de p_0, la propriété locale de l'application D nous dit qu'au voisinage de p_0 on a $Dg_1 = Dg_1^*$, donc en particulier

$$(Dg_1)\,(p_0) = (Dg_1^*)\,(p_0). \tag{19}$$

Donc, le second membre de la formule (17) est indépendant du choix de la fonction g_1 et est défini exclusivement par la fonction g et le

point p_0. Ceci exprime que sur U la fonction $D_U g$ dépend uniquement de g. Donc, l'application $D_U: g \mapsto D_U g$ est unique.

Pour prouver son existence on prendra la formule (18) pour définition de la fonction $D_U g$. Cette définition est cohérente d'après la formule (19). Bien plus, si $g_1 = g$ dans un voisinage W de p_0, alors pour tout point $p \in W$ la fonction $g_1 = g_{1, p_0}$ peut tenir le rôle de la fonction $g_{1, p}$ dans le calcul de la valeur $(D_U g)(p)$ de $D_U g$ au point p. Ceci exprime que $D_U g = Dg_1$ non seulement en p_0 mais à proximité de p_0. Donc, à proximité de chaque point de U, la fonction $D_U g$ est confondue avec une fonction différentiable sur $\mathscr{X}$. Par conséquent, la fonction $D_U g$ est différentiable sur U, i.e. la correspondance $g \mapsto D_U g$ est l'application

$$D_U: \mathsf{F}U \to \mathsf{F}U.$$

Si f, g sont des fonctions différentiables sur U, et f_1, g_1 des fonctions différentiables sur $\mathscr{X}$, respectivement confondues avec f et g à proximité de p_0, la fonction $f_1 + g_1$ sera confondue avec $f + g$ à proximité de p_0. Donc,

$$[D_U(fg)](p_0) = [D(f_1 g_1)](p_0) = \\ = [Df_1 \cdot g_1 + f_1 \cdot Dg_1](p_0) = [D_U f \cdot g + f \cdot D_g](p_0),$$

et, par suite, l'application D_U, qui est visiblement linéaire, est une dérivation sur U.

Enfin, si f est une fonction différentiable sur $\mathscr{X}$, pour la fonction $g = f|_U$ le rôle de la fonction g_1 (pour tout point $p_0 \in U$) peut être tenu par la fonction f. Donc,

$$D_U(f|_U) = Df|_U. \quad \square$$

D'une façon générale, au lieu de $D_U g$ on écrira tout simplement Dg.

Pour démontrer le théorème 1, nous aurons encore besoin du lemme suivant:

Lemme 2. *Soit $(U, h) = (U, x^1, \ldots, x^n)$ une carte d'une variété $\mathscr{X}$ de classe C^∞. Au voisinage de tout point $p_0 \in U$, chaque fonction f différentiable (sur U) se représente sous la forme*

$$f = f(p_0) + (x^i - x_0^i) f_i, \tag{20}$$

où $f_1, \ldots, f_n$ sont des fonctions différentiables (au voisinage de p_0) et $x_0^1, \ldots, x_0^n$ les coordonnées de p_0.

Démonstration. Soit $\mathbf{x}_0 = h(p)$ et soit $r > 0$ un nombre tel que tout point $\mathbf{x} \in \mathbb{R}^n$ tel que $|\mathbf{x} - \mathbf{x}_0| < r$ appartienne à l'ensemble ouvert $h(U)$. Alors pour la valeur $f(\mathbf{x})$, en un tel point $\mathbf{x}$,

d'une fonction arbitraire f différentiable sur $h(U)$, on aura

$$f(\mathbf{x})-f(\mathbf{x}_0)=\int_0^1 \frac{\partial}{\partial s} f(\mathbf{x}_0+s(\mathbf{x}-\mathbf{x}_0))\,\mathrm{d}s=$$

$$=(x^i-x_0^i)\int_0^1 \frac{\partial f}{\partial x^i}(\mathbf{x}_0+s(\mathbf{x}-\mathbf{x}_0))\,\mathrm{d}s,$$

i.e. la formule

$$f(\mathbf{x})=f(\mathbf{x}_0)+(x^i-x_0^i)f_i(\mathbf{x}),$$

où

$$f_i(\mathbf{x})=\int_0^1 \frac{\partial f}{\partial x^i}(\mathbf{x}_0+s(\mathbf{x}-\mathbf{x}_0))\,\mathrm{d}s.$$

Pour achever la démonstration, il reste à passer à l'aide du difféomorphisme h des fonctions sur $h(U)$ aux fonctions sur U. □

Pour toute dérivation D d'une algèbre $\mathcal{A}$ à unité 1, on a l'égalité

$$D1 = D(1\cdot 1) = D1\cdot 1 + 1\cdot D1 = D1 + D1 = 2D1,$$

donc,

$$D1 = 0.$$

L'application D étant linéaire, on en déduit que $Da = 0$ *pour tout élément* $a \in \mathbb{R}$ *du champ principal.*

Pour les dérivations sur une variété différentiable $\mathcal{X}$ (i.e. les dérivations de l'algèbre $\mathsf{F}\mathcal{X}$) ceci exprime que *toute dérivation sur* $\mathcal{X}$ *donne un zéro d'une fonction constante.*

Nous disposons maintenant de tous les éléments pour prouver le théorème 1.

D é m o n s t r a t i o n d u t h é o r è m e 1. Comme toujours on commencera par établir l'unicité.

Supposons qu'une dérivation D sur $\mathcal{X}$ est engendrée par un champ de vecteurs X. Ceci exprime que dans chaque carte $(U, x^1, \ldots, x^n)$ de la variété $\mathcal{X}$, toute fonction f différentiable sur U est justiciable de l'égalité

$$Df = X^i \frac{\partial f}{\partial x^i} \text{ sur } U,$$

où X^i, $i = 1, \ldots, n$, sont les composantes du champ de vecteurs X sur U (et où le symbole D désigne en fait D_U). En particulier,

$$Dx^i = X^i, \quad i = 1, \ldots, n,$$

ce qui prouve l'unicité de X.

Supposons maintenant que D est une dérivation arbitraire sur $\mathscr{X}$ (qui, rappelons-le, est supposée séparée et de classe C^∞). Pour toute carte $(U, x^1, \ldots, x^n)$, considérons les fonctions différentiables sur U

$$X^i = Dx^i, \quad i = 1, \ldots, n.$$

Soit f une fonction différentiable sur U. En appliquant l'opérateur D (ou plus exactement l'opérateur D_W, où W est un voisinage du point p_0 en lequel a lieu (19)) à sa représentation (20), on trouve que

$$Df = X^i f_i + (x^i - x_0^i)\, Df_i \quad \text{sur } W,$$

donc que

$$(Df)(p_0) = X^i(p_0)\, f_i(p_0). \tag{21}$$

(La propriété locale de D nous permet de remplacer D_W par D.)

L'égalité (21) est évidemment valable non seulement pour les dérivations sur $\mathscr{X}$, mais aussi pour toute dérivation sur U. En particulier, elle l'est pour la dérivation $\frac{\partial}{\partial x^i}$. Donc,

$$\left(\frac{\partial}{\partial x^i} f\right)(p_0) = f_i(p_0),$$

et, par suite, en vertu de la formule (21)

$$(Df)(p_0) = X^i(p_0)\frac{\partial f}{\partial x^i}(p_0) = \left(X^i \frac{\partial f}{\partial x^i}\right)(p_0).$$

Le point p_0 étant un point arbitraire du voisinage U, ceci prouve que

$$Df = X^i \frac{\partial f}{\partial x^i} \text{ sur } U,$$

i.e. la dérivation D est engendrée sur U par le champ de vecteurs X_U de composantes $X^1, \ldots, X^n$.

De façon analogue, pour toute autre carte $(U', x^{1'}, \ldots, x^{n'})$ la dérivation D sur U' est engendrée par le champ de vecteurs $X_{U'}$ de composantes $X^{i'} = Dx^{i'}$. Ceci étant, d'après la propriété locale de l'opérateur D, pour toute fonction f on aura sur l'intersection $U \cap U'$

$$X^{i'} \frac{\partial f}{\partial x^{i'}} = X^i \frac{\partial f}{\partial x^i}.$$

En appliquant cette égalité aux fonctions de passage $x^{i'} = x^{i'}(x^1, \ldots, x^n)$, on obtient immédiatement

$$X^{i'} = X^i \frac{\partial x^{i'}}{\partial x^i},$$

i.e. les champs de vecteurs X_U et $X_{U'}$ sont confondus sur $U \cap U'$. Donc, en posant pour tout point $p \in \mathscr{X}$

$$X_p = (X_U)_p,$$

où U est un voisinage de coordonnées de p, on définit de façon unique un champ de vecteurs X sur $\mathscr{X}$, tel que $X|_U = X_U$ quel que soit U. Donc, ce champ est *primo* différentiable, *secundo* engendre la dérivation D. □

Question de contrôle. A quel endroit de la démonstration a-t-on utilisé l'hypothèse que $\mathscr{X}$ est une variété de classe C^∞ ?

En vertu du théorème 1, les champs de vecteurs sur $\mathscr{X}$ sont généralement identifiés aux dérivations. (Ce qui en particulier justifie *a posteriori* la notation Xf pour représenter l'action sur f de la dérivation engendrée par le champ de vecteurs X.)

* * *

Il est clair que pour toute algèbre $\mathscr{A}$, la somme de deux dérivations ainsi que le produit d'une dérivation par un nombre sont des dérivations, i.e. *l'ensemble* Der $\mathscr{A}$ *de toutes les dérivations de l'algèbre* $\mathscr{A}$ *est un espace vectoriel* (qui est un sous-espace de l'espace vectoriel $\mathrm{End}_{\mathrm{lin}}\mathscr{A}$ de tous les opérateurs linéaires $\mathscr{A} \to \mathscr{A}$).

On appelle *commutateur* (ou *crochet de Lie*) de deux opérateurs linéaires $D_1, D_2 : \mathscr{A} \to \mathscr{A}$ l'opérateur

$$[D_1, D_2] = D_1D_2 - D_2D_1.$$

Il est aisé de voir que *pour toute algèbre* $\mathscr{A}$ *le commutateur* $[D_1, D_2]$ *de deux dérivations arbitraires* $D_1, D_2 : \mathscr{A} \to \mathscr{A}$ *est aussi une dérivation*. En effet,

$$\begin{aligned}
[D_1, D_2]\,(ab) &= D_1\,(D_2\,(ab)) - D_2\,(D_1\,(ab)) = \\
&= D_1\,(D_2a\cdot b + a\cdot D_2b) - D_2\,(D_1a\cdot b + a\cdot D_1b) = \\
&= D_1D_2a\cdot b + D_2a\cdot D_1b + D_1a\cdot D_2b + a\cdot D_1D_2b - \\
&\quad - D_2D_1a\cdot b - D_1a\cdot D_2b - D_2a\cdot D_1b - a\cdot D_2D_1b = \\
&= (D_1D_2 - D_2D_1)\,a\cdot b + a\cdot(D_1D_2 - D_2D_1)\,b = \\
&= [D_1, D_2]\,a\cdot b + a\cdot\,[D_1, D_2)\,b
\end{aligned}$$

pour des éléments arbitraires $a, b \in \mathscr{A}$. □

L'opération de commutation $D_1, D_2 \mapsto [D_1, D_2]$ étant visiblement linéaire par rapport à D_1 et D_2, ceci exprime que *l'espace vectoriel* Der $\mathscr{A}$ *des dérivations de l'algèbre* $\mathscr{A}$ *est une algèbre par rappart à la commutation*.

Il est clair que *la commutation est anticommutative*, i.e.

$$[D_1, D_2] = -[D_2, D_1]$$

pour des opérateurs linéaires quelconques D_1 et D_2 (qui ne sont pas nécessairement des dérivations).

En outre, *pour tout triplet d'opérateurs* D_1, D_2, D_3, *on a l'identité*

$$[[D_1, D_2], D_3] + [[D_2, D_3], D_1] + [[D_3, D_1], D_2] = 0. \tag{22}$$

En effet,

$$[[D_1, D_2], D_3] = (D_1D_2 - D_2D_1)\, D_3 - D_3\,(D_1D_2 - D_2D_1) = \\ = D_1D_2D_3 - D_2D_1D_3 - D_3D_1D_2 + D_3D_2D_1,$$

donc, le premier membre de la formule (22) est la somme de douze opérateurs de la forme $D_iD_jD_k$, dans laquelle chacun d'eux figure deux fois avec des signes contraires. Cette somme est donc nulle. □

L'identité (22) s'appelle *identité de Jacobi.*

Ceci nous suggère la définition générale suivante:

Définition 2. On appelle *algèbre de Lie* une algèbre munie d'une multiplication anticommutative et vérifiant l'identité de Jacobi.

On a ainsi prouvé que *l'algèbre* Der $\mathcal{A}$ *est une algèbre de Lie.*

Les dérivations de l'algèbre des fonctions différentiables $\mathsf{F}\mathcal{X}$ n'étant, en vertu des identifications effectuées, autres que des champs de vecteurs sur $\mathcal{X}$, on trouve en particulier que *les champs de vecteurs sur une variété séparée différentiable* $\mathcal{X}$ *de classe* C^∞ *forment une algèbre de Lie.*

Cette algèbre est généralement désignée par $\mathfrak{a}\mathcal{X}$.

Par définition,

$$[X, Y]\, f = X\,(Yf) - Y\,(Xf)$$

pour tous champs de vecteurs X, Y sur $\mathcal{X}$ et toute fonction $f \in \mathsf{F}\mathcal{X}$.

Il s'ensuit que

$$[gX, Y] = g\,[X, Y] - Yg\cdot X \tag{23}$$

pour toute fonction $g \in \mathsf{F}\mathcal{X}$. En effet,

$$[gX, Y]\, f = gX\,(Yf) - Y\,(gXf) = \\ = gX\,(Yf) - Yg\cdot Xf - gY\,(Xf) = g\,[X, Y]\, f - Yg\cdot Xf$$

pour toute fonction $f \in \mathsf{F}\mathcal{X}$. □

Si

$$X = X^i \frac{\partial}{\partial x^i} \quad \text{et} \quad Y = Y^j \frac{\partial}{\partial x^j},$$

dans une carte $(U, x^1, \ldots, x^n)$, on a pour toute fonction f différentiable sur U

$$[X, Y]\, f = X\,(Yf) - Y\,(Xf) = X\left(Y^j \frac{\partial f}{\partial x^j}\right) - Y\left(X^i \frac{\partial f}{\partial x^i}\right) = \\ = X^i \frac{\partial}{\partial x^i}\left(Y^j \frac{\partial f}{\partial x^j}\right) - Y^j \frac{\partial}{\partial x^j}\left(X^i \frac{\partial f}{\partial x^i}\right) =$$

$$= X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial f}{\partial x^j} + X^i Y^j \frac{\partial^2 f}{\partial x^i \partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \frac{\partial f}{\partial x^i} - Y^j X^i \frac{\partial^2 f}{\partial x^i \partial x^j} =$$
$$= \left(X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \right) \frac{\partial f}{\partial x^i}.$$

Ceci prouve que *dans chaque carte* $(U, x^1, \ldots, x^n)$ *les composantes* $[X, Y]^i$ *du champ de vecteurs* $[X, Y]$ *s'expriment par la formule*

$$(24) \qquad [X, Y]^i = X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j}, \quad i = 1, \ldots, n,$$

où $X^1, \ldots, X^n, Y^1, \ldots, Y^n$ *sont les composantes des champs* X *et* Y *respectivement.*

Remarque 4. Nous avons déjà utilisé le terme « algèbre de Lie » pour les espaces vectoriels de matrices envisagées par la définition des groupes de Lie de matrices (cf. définition 1 de la leçon 11). Il est évident que l'espace vectoriel $\mathrm{Mat}_n(\mathbb{R}) = \mathfrak{gl}_n(\mathbb{R})$ des matrices d'ordre n est une algèbre de Lie par rapport à la commutation des matrices $[B_1, B_2] = B_1 B_2 - B_2 B_1$. Donc, il en sera de même de toute sous-algèbre (de tout sous-espace vectoriel $\mathfrak{g}$ stable pour cette opération). D'autre part, si $\mathfrak{g}$ est le sous-espace de la définition 1 de la leçon 11 et si $B_1, B_2 \in \mathfrak{g}$, les matrices e^{tB_1} et e^{tB_2}, $t \in \mathbb{R}$, appartiennent au groupe $\mathcal{G}$ en vertu de la condition 1° de cette définition. Donc, ce groupe contiendra aussi la matrice

$$A_t = e^{tB_1} e^{tB_2} e^{-tB_1} e^{-tB_2} =$$
$$= (E + B_1 t + \ldots)(E + B_2 t + \ldots)(E - B_1 t + \ldots)(E - B_2 t + \ldots) =$$
$$= E + [B_1, B_2] t^2 + \ldots,$$

où les points de suspension désignent les termes de puissance supérieure à 3 en t. Puisque $A_t \to E$ lorsque $t \to 0$, la norme de la matrice $A_t - E$ est < 1 pour t assez petit et, par suite, est définie la matrice

$$B_t = \ln A_t = (A_t - E) - \frac{(A-E)^2}{2} + \cdots = [B_1, B_2] t^2 + \ldots$$

Comme $B_t \to 0$ lorsque $t \to 0$, sa norme $|B_t|$ vérifie l'inégalité $|B_t| < \ln 2$ pour t assez petit et, par conséquent, la matrice B_t appartient à $\mathfrak{g}$ d'après la condition 2° de la définition 1 donnée à la leçon 11. Ce sous-espace contiendra alors la matrice $\frac{B_t}{t^2}$, donc la matrice

$$\lim_{t \to 0} \frac{B_t}{t^2} = \lim_{t \to 0} ([B_1, B_2] + \ldots) = [B_1, B_2].$$

Par conséquent, $\mathfrak{g}$ est une algèbre de Lie (une sous-algèbre de l'algèbre de Lie $\mathfrak{gl}_n(\mathbb{R})$). Les sous-algèbres des algèbres de Lie $\mathfrak{gl}_n(\mathbb{R})$ s'appellent *algèbres de Lie de matrices.*

Remarque 5. A noter qu'il existe des algèbres de Lie de matrices qui ne sont des algèbres de Lie d'aucun groupe de Lie de matrices. Nous reviendrons sur cette question au prochain tome.

LEÇON 17

Courbes intégrales de champs de vecteurs. — Champs de vecteurs et flots. — Transfert de champs de tenseurs par des difféomorphismes. — Dérivée de Lie d'un champ de tenseurs.

On appelle *courbe* sur une variété différentiable $\mathscr{X}$ une application différentiable

$$\gamma :]a, b[\to \mathscr{X} \tag{1}$$

d'un intervalle $]a, b[$ de $\mathbb{R}$ dans $\mathscr{X}$ (cf. notion de courbe à la leçon 1).

La différentielle $(d\gamma)_t$ d'une courbe γ en tout point $t \in]a, b[$ est une application linéaire de l'espace à une dimension $T_t\,]a, b[= \mathbb{R}$ dans l'espace $T_{\gamma(t)}\mathscr{X}$, caractérisée de façon unique par le vecteur

$$\dot{\gamma}(t) = (d\gamma)_t \left(\frac{\partial}{\partial t}\right)_t \tag{2}$$

de l'espace $T_{\gamma(t)}\mathscr{X}$ en lequel elle transforme le vecteur de base $\left(\frac{\partial}{\partial t}\right)_t$ de l'espace $T_t\,]a, b[$.

Définition 1. Le vecteur $\dot{\gamma}(t)$ s'appelle *vecteur tangent* à la courbe γ en t. (Par abus de langage, le vecteur $\dot{\gamma}(t)$ s'appelle aussi vecteur tangent à la courbe γ au point $p = \gamma(t)$.)

Soit maintenant X un champ de vecteurs sur $\mathscr{X}$. Par définition, le champ X associe à chaque point $p \in \mathscr{X}$ un vecteur $X_p \in T_p\mathscr{X}$.

Définition 2. Une courbe γ s'appelle *courbe intégrale* (ou *trajectoire*) d'un champ de vecteurs X si

$$\dot{\gamma}(t) = X_{\gamma(t)} \quad \text{pour tout } t \in]a, b[. \tag{3}$$

On dit qu'une courbe γ est *contenue* dans une carte $(U, h) = (U, x^1, \ldots, x^n)$ si $\gamma(t) \in U$ pour tout $t \in]a, b[$. Une telle courbe est définie par n fonctions différentiables

$$x^i = x^i(t), \quad a < t < b, \quad i = 1, \ldots, n, \tag{4}$$

et le vecteur tangent $\dot{\gamma}(t)$ vérifie l'égalité

$$\dot{\gamma}(t) = \dot{x}^i(t) \left(\frac{\partial}{\partial x^i}\right)_{\gamma(t)}$$

(prouvez-le !). Donc, pour une courbe dans U, l'équation (3) équivaut à un système de n équations différentielles

$$\dot{x}^i(t) = X^i(x^1(t), \ldots, x^n(t)), \quad i = 1, \ldots, n, \tag{5}$$

du premier ordre, où X^i, $i = 1, \ldots, n$, sont les composantes du champ de vecteurs X dans la carte (U, h) (ou, plus exactement, leurs expressions en fonction des coordonnées $x^1, \ldots, x^n$).

On voit donc que les équations (3) généralisent, à des variétés arbitraires, la notion de système d'équations différentielles du premier ordre défini dans un domaine de l'espace $\mathbb{R}^n$. On les appelle *équations différentielles sur la variété* $\mathcal{X}$. Ces équations sont étudiées en théorie des équations différentielles et sortent pour l'essentiel du cadre de notre exposé.

Toutefois, pour compléter notre exposé et mettre en évidence le trait spécifique des équations générales (3), nous allons formuler et prouver le théorème fondamental d'existence et d'unicité de leurs solutions.

Il est clair que pour tout intervalle partiel I' de l'intervalle $I =]a, b[$ la restriction $\gamma|_{I'}$ de la courbe intégrale γ à I' sera aussi une courbe intégrale du champ de vecteurs X. Une courbe intégrale (1) est dite *maximale* si elle n'est la restriction d'aucune courbe intégrale définie sur un plus grand intervalle.

Soit $t_0 \in \mathbb{R}$. On dit qu'une courbe (1) *passe pour* $t = t_0$ *par un point* $p \in \mathcal{X}$ si *primo* elle est définie sur un intervalle $]a, b[$ de $\mathbb{R}$ tel que $a < t_0 < b$ et *secundo* $\gamma(t_0) = p$.

Théorème 1. *Si une variété* $\mathcal{X}$ *est séparée, pour tout point* $p_0 \in \mathcal{X}$ *et tout champ de vecteurs* X *sur* $\mathcal{X}$, *il existe une seule courbe intégrale maximale* $\gamma: I \to \mathcal{X}$ *du champ* X *passant par* p_0 *pour* $t = t_0$.

Démonstration. Soit Γ l'ensemble de toutes les courbes intégrales γ du champ de vecteurs X passant pour $t = t_0$ par le point p_0. L'équation (3) étant équivalente à un système (5) d'équations différentielles ordinaires dans toute carte (U, h) contenant le point p_0, le théorème classique d'existence et d'unicité de la solution de tels systèmes affirme l'existence d'au moins une courbe intégrale γ du champ X telle que $\gamma(t_0) = p_0$. Ceci exprime que *l'ensemble* Γ *n'est pas vide.*

Soient maintenant $\gamma_1: I_1 \to \mathcal{X}$ et $\gamma_2: I_2 \to \mathcal{X}$ deux courbes intégrales de Γ. Considérons l'intersection $I_0 = I_1 \cap I_2$ des intervalles I_1 et I_2, qui est de toute évidence un intervalle de l'axe $\mathbb{R}$ contenant le point p_0, et son sous-ensemble C composé de tous les points $t \in I_0$ tels que $\gamma_1(t) = \gamma_2(t)$. Par hypothèse, $t_0 \in C$, donc *l'ensemble* C *n'est pas vide.*

Soit $t_1 \in C$ et soit (U, h) une carte contenant le point $p_1 = \gamma_1(t_1) = \gamma_2(t_1)$. En vertu du théorème d'existence et d'unicité des solutions de systèmes d'équations différentielles, sur un intervalle I de $\mathbb{R}$ contenant le point t_1, il existe une seule courbe inté-

grale $\gamma: I \to \mathscr{X}$ du champ X telle que $\gamma(t_1) = p_1$. Ceci étant, on peut admettre sans perdre en généralité que $I \subset I_0$. Dans ces conditions, les restrictions $\gamma_1|_I$ et $\gamma_2|_I$ seront une telle courbe γ et, en vertu de l'unicité, seront confondues. Par définition, ceci exprime que $I \subset C$. Donc, pour tout point $t_1 \in C$ il existe un intervalle I contenant ce point, tel que $I \subset C$. Donc, *l'ensemble C est ouvert* (dans $\mathbb{R}$ et, par suite, dans I_0).

Considérons maintenant une application $\gamma_1 \times \gamma_2: I_0 \to \mathscr{X} \times \mathscr{X}$ définie par la formule

$$(\gamma_1 \times \gamma_2)(t) = (\gamma_1(t), \gamma_2(t)), \quad t \in I.$$

Il est clair que l'application $\gamma_1 \times \gamma_2$ est continue et que l'ensemble C n'est autre que l'image réciproque $(\gamma_1 \times \gamma_2)^{-1}\Delta$ par cette application de la diagonale $\Delta = \{(p, p); p \in \mathscr{X}\}$ du produit $\mathscr{X} \times \mathscr{X}$ (cf. leçon 15). La diagonale Δ étant fermée dans le cas d'une variété $\mathscr{X}$ séparée, on en déduit que *l'ensemble C est fermée dans I_0*.

L'intervalle I_0 étant connexe, on a prouvé que $C = I_0$. Ceci exprime que *deux courbes intégrales quelconques de Γ sont confondues sur l'intersection de leurs domaines de définition.*

Il s'ensuit que si l'on désigne par I_γ l'intervalle sur lequel est définie la courbe $\gamma \in \Gamma$ et si l'on pose

$$I = \bigcup_{\gamma \in \Gamma} I_\gamma$$

et

$$\gamma_0(t) = \gamma(t), \quad t \in I_\gamma,$$

on définit de façon cohérente l'unique courbe intégrale maximale $\gamma_0: I \to \mathscr{X}$ du champ X pour laquelle $\gamma(t_0) = p_0$. □

Exercice 1. Construire une variété non séparée à deux dimensions $\mathscr{X}$ et un champ de vecteurs X sur $\mathscr{X}$, tels que pour un point $p_0 \in \mathscr{X}$ il existe deux courbes intégrales du champ X passant par p_0 pour $t = t_0$.

* * *

Soit $t_0 = 0$.

Pour tout point $p \in \mathscr{X}$ désignons par γ_p^X la courbe intégrale maximale du champ X passant par p pour $t = 0$, et pour tout $t \in \mathbb{R}$ pour lequel le point $\gamma_p^X(t)$ est défini posons

$$\varphi_t^X(p) = \gamma_p^X(t).$$

Donc, φ_t^X est une application dans $\mathscr{X}$ du sous-ensemble $D_t \subset \mathscr{X}$ composé des points $p \in \mathscr{X}$ pour lesquels est défini le point $\gamma_t^X(t)$.

Exercice 2. Montrer que *l'ensemble D_t est ouvert (dans $\mathscr{X}$) et que (pour $D_t \neq \varnothing$) l'application $\varphi_t^X: D_t \to \mathscr{X}$ est différentiable.* [N o t a. Se servir du théorème de dépendance des solutions de systèmes d'équations différentielles par rapport aux conditions initiales.]

Exercice 3. Montrer que les applications $\varphi_t = \varphi_t^X$ jouissent des propriétés suivantes:

a) Il existe une fonction continue $\varepsilon: \mathcal{X} \to \mathbb{R}$ prenant des valeurs strictement positives, telle que $p \in D_t$ pour $|t| < \varepsilon(p)$.

b) L'application φ_0 est définie sur $\mathcal{X}$ tout entière (i.e. $D_0 = \mathcal{X}$) et est l'application identique id de $\mathcal{X}$.

c) Si $\varphi_t(p) \in D_s$ (en particulier, si $|s| < \varepsilon(\varphi_t(p))$), alors $p \in D_{s+t}$ et

$$\varphi_s(\varphi_t(p)) = \varphi_{s+t}(p). \tag{6}$$

[N o t a. Pour prouver la proposition a) se servir du théorème de dépendance des solutions des équations différentielles par rapport aux données initiales. La proposition b) est évidente. Pour établir la proposition c) il suffit de remarquer que les courbes $s \mapsto \gamma^X_{\varphi_t(p)}(s)$ et $s \mapsto \gamma^X_p(s+t)$ sont des courbes intégrales maximales du champ X, passant par le point $\varphi_t(p)$ pour $s = 0$.]

Par abus de langage, la propriété c) est généralement exprimée par l'identité

$$\varphi_s \circ \varphi_t = \varphi_{s+t}.$$

Définition 3. Une famille d'applications différentiables $\varphi_t: D_t \to \mathcal{X}$ jouissant des propriétés a), b) et c) s'appelle *flot* sur la variété $\mathcal{X}$.

On voit donc que *chaque champ de vecteurs $X \in \mathfrak{a}\mathcal{X}$ induit un flot $\{\varphi_t^X\}$ sur $\mathcal{X}$.*

Réciproquement, tout flot $\{\varphi_t\}$ définit un champ de vecteurs X sur $\mathcal{X}$ à l'aide de la formule

$$X_p = \dot{\gamma}_p(0), \quad p \in \mathcal{X},$$

où γ_p est une courbe $t \mapsto \varphi_t(p)$, $|t| < \varepsilon(p)$.

Un flot $\{\varphi'_t: D'_t \to \mathcal{X}\}$ s'appelle *partie* d'un flot $\{\varphi_t: D_t \to \mathcal{X}\}$ si $D'_t \subset D_t$ et $\varphi_t|_{D'_t} = \varphi'_t$ pour tout $t \in \mathbb{R}$. Un flot $\{\varphi_t\}$ est dit *maximal* s'il n'est la partie d'aucun autre flot.

Il est clair que

1° Un flot $\{\varphi_t\}$ et toute partie $\{\varphi'_t\}$ de ce flot engendrent le même champ de vecteurs X.

2° Le flot $\{\varphi_t^X\}$ induit par le champ de vecteurs X est maximal.

Donc, *la formule*

$$\textit{champ } X \Rightarrow \textit{flot } \{\varphi_t^X\}$$

établit une correspondance biunivoque entre les champs de vecteurs et les flots maximaux sur $\mathcal{X}$.

La fonction ε étant continue et donc

$$\lim_{t \to 0} \varepsilon(\varphi_t(p)) = \varepsilon(p) > 0$$

pour tout point $p \in \mathscr{X}$, il existe une fonction continue $\delta: \mathscr{X} \to \mathbb{R}$ telle que

$$|t| < \varepsilon(\varphi_t(p)) \quad \text{pour} \quad |t| < \delta(p).$$

Soit O_t un sous-ensemble ouvert de la variété $\mathscr{X}$, composé de tous les points $p \in \mathscr{X}$ tels que $|t| < \delta(p)$. Il existe alors un $\delta_0 > 0$, plus exactement $\delta_0 = \delta(p_0)$, tel que $p_0 \in O_t$ pour $|t| < \delta_0$ (de sorte que l'ensemble O_t n'est visiblement pas vide pour t assez petit).

Puisque pour $p \in O_t$, i.e. pour $|t| < \delta(p)$, est défini un point $\varphi_{-t}(\varphi_t(p))$ confondu, en vertu de la propriété c), avec le point p, la restriction de l'application φ_t à l'ensemble O_t est une application bijective de cet ensemble sur l'ensemble (visiblement ouvert) $O'_t = \varphi_t O_t$ (l'application réciproque est $\varphi_{-t}|_{O'_t}$). Les applications φ_t et φ_{-t} étant différentiables par définition, ceci prouve que *pour tout $t \in \mathbb{R}$ (pour lequel l'ensemble O_t n'est pas vide) l'application φ_t est un difféomorphisme de O_t sur O'_t.*

* * *

Soit maintenant $\varphi: \mathscr{X} \to \mathscr{Y}$ un difféomorphisme de variétés différentiables et soit S un champ de tenseurs de type (a, b) sur une variété $\mathscr{Y}$. Par définition, les composantes $S^{j_1 \ldots j_b}_{i_1 \ldots i_a}$ de S dans chaque carte (V, k) sur $\mathscr{Y}$ sont des fonctions différentiables sur V. D'autre part, pour toute carte (U, h) sur $\mathscr{X}$ le couple (V, k), où $V = \varphi U$ et $k = h \circ \varphi^{-1}$, sera visiblement une carte sur $\mathscr{Y}$. A la limite de ce qui vient d'être dit, définissons un champ de tenseurs $\varphi^* S$ sur $\mathscr{X}$ en admettant que dans chaque carte (U, h) ses composantes sont

$$(\varphi^* S)^{j_1 \ldots j_b}_{i_1 \ldots i_a} = S^{j_1 \ldots j_b}_{i_1 \ldots i_a} \circ (\varphi|_U), \tag{7}$$

où $S^{j_1 \ldots j_b}_{i_1 \ldots i_a}$ sont les composantes du champ S dans la carte $(\varphi U, h \circ \varphi^{-1})$. Etant donné que pour deux cartes quelconques (U, h) et (U', h') sur $\mathscr{X}$, le changement de cartes $h' \circ h^{-1}$ est confondu avec $(h' \circ \varphi^{-1}) \circ (h \circ \varphi^{-1})^{-1}$ pour les cartes $(\varphi U, h \circ \varphi^{-1})$ et $(\varphi U', h' \circ \varphi^{-1})$, les fonctions $(\varphi^* S)^{j_1 \ldots j_b}_{i_1 \ldots i_a}$ dans diverses cartes seront liées (dans l'intersection de ces cartes) par la même loi tensorielle de transformation que les composantes $S^{j_1 \ldots j_b}_{i_1 \ldots i_a}$ du champ S. Donc, ces fonctions seront bien les composantes d'un champ de tenseurs $\varphi^* S$.

On dira du champ $\varphi^* S$ qu'il est le résultat du *transfert* du champ S de $\mathscr{Y}$ sur $\mathscr{X}$ par le difféomorphisme φ.

Exemple 1. Si le champ S est de type (0, 0) (est une fonction différentiable f sur $\mathscr{Y}$), alors

$$\varphi^* f = f \circ \varphi. \tag{8}$$

Exemple 2. Si le champ S est de type (0, 1) (est un champ de vecteurs X), alors

$$(\varphi^* X)_p = (\mathrm{d}\varphi_p)^{-1} X_{\varphi(p)} \tag{9}$$

pour tout point $p \in \mathscr{X}$, où $(\mathrm{d}\varphi_p)^{-1} : \mathrm{T}_{\varphi(p)}\mathscr{Y} \to \mathrm{T}_p\mathscr{X}$ est l'application réciproque de l'isomorphisme $\mathrm{d}\varphi_p : \mathrm{T}_p\mathscr{X} \to \mathrm{T}_{\varphi(p)}\mathscr{Y}$.

Exemple 3. Si le champ S est de type (1, 0) (est un champ de covecteurs α), alors

$$(\varphi^* \alpha)_p = (\mathrm{d}\varphi_p)^* \alpha_{\varphi(p)} \tag{10}$$

pour tout point $p \in \mathscr{X}$, où $(\mathrm{d}\varphi_p)^*$ est l'application $\mathrm{T}^*_{\varphi(p)}\mathscr{Y} \to \mathrm{T}^*_p\mathscr{X}$ adjointe de $\mathrm{d}\varphi_p : \mathrm{T}_p\mathscr{X} \to \mathrm{T}_{\varphi(p)}\mathscr{Y}$.

Exercice 4. Décrire de façon analogue les tenseurs $(\varphi^* S)_p$ d'un champ de tenseurs S de type quelconque (a, b). [N o t a. La correspondance $S \mapsto \varphi^* S$ préserve toutes les opérations algébriques sur les champs de tenseurs. En particulier,

$$\varphi^* (S \otimes T) = \varphi^* S \otimes \varphi^* T \tag{11}$$

pour tous champs de tenseurs S et T sur $\mathscr{X}$.]

Remarque 1. Il est aisé de voir que *pour tous champs de vecteurs* $X, Y \in \mathfrak{a}\mathscr{Y}$ *et tout difféomorphisme* $\varphi : \mathscr{X} \to \mathscr{Y}$ *on a*

$$[\varphi^* X, \varphi^* Y] = \varphi^* [X, Y],$$

i.e. l'application $\varphi^* : \mathfrak{a}\mathscr{Y} \to \mathfrak{a}\mathscr{X}$ (visiblement linéaire) est un isomorphisme de l'algèbre de Lie $\mathfrak{a}\mathscr{Y}$ sur l'algèbre de Lie $\mathfrak{a}\mathscr{X}$. (Pour le prouver il suffit de remarquer que les deux membres de cette formule coïncident visiblement en coordonnées locales, coordonnées sur lesquelles φ agit par égalisation.)

Remarque 2. On voit immédiatement sur la formule (9) que le transfert du champ de vecteurs n'est généralement possible que par un difféomorphisme. Au contraire, la formule (10) a un sens pour toute application différentiable $\varphi : \mathscr{X} \to \mathscr{Y}$. Donc, les champs de covecteurs peuvent être transférés par des applications quelconques. Nous reviendrons sur ce sujet à la prochaine leçon.

Remarque 3. En désignant le champ X par Y, et le champ $\varphi^* X$ par X, on peut mettre la formule (9) sous la forme suivante

$$Y_{\varphi(p)} = (\mathrm{d}\varphi)_p X_p. \tag{9'}$$

Sous cette forme elle a un sens pour toute application différentiable $\varphi : \mathscr{X} \to \mathscr{Y}$. Des champs $X \in \mathfrak{a}\mathscr{X}$, $Y \in \mathfrak{a}\mathscr{Y}$ vérifiant l'égalité (9′) en chaque point $p \in \mathscr{X}$ sont dits φ-*liés*.

Exercice 5. Montrer que *si des champs* $X_1, X_2 \in \mathfrak{a}\mathscr{X}$ *sont* φ-*liés respectivement avec des champs* $Y_1, Y_2 \in \mathfrak{a}\mathscr{Y}$, *le champ* $[X_1, X_2]$ *le sera avec* $[Y_1, Y_2]$. [N o t a. *Des champs* $X \in \mathfrak{a}\mathscr{X}$ et $Y \in \mathfrak{a}\mathscr{Y}$ *sont* φ-*liés si et seulement si pour toute fonction* $f \in \mathrm{F}\mathscr{Y}$ *on a* $X(f \circ \varphi) = Yf \circ \varphi$.]

* * *

Appliquons maintenant la construction générale exposée au cas particulier de difféomorphismes de la forme φ_t.

Soit S un champ de tenseurs de type (a, b) sur une variété $\mathscr{X}$ et soit $p \in \mathscr{X}$.

Par définition, $p \in O_t$ pour $|t| < \delta(p)$, donc pour tout t tel que $|t| < \delta(p)$, est défini au point p le tenseur $(\varphi_t^* S)_p$ (où S désigne de toute évidence la restriction de S à O_t), et, par suite, le tenseur $(\varphi_t^* S)_p - S_p$. Posons

$$(\pounds_X S)_p = \lim_{t \to 0} \frac{(\varphi_t^* S)_p - S_p}{t}, \tag{12}$$

où X est le champ de vecteurs engendré par le flot $\{\varphi_t\}$.

Le point p étant un point arbitraire de la variété $\mathscr{X}$, les tenseurs (12) forment un champ de tenseurs $\pounds_X S$ de type (a, b) sur $\mathscr{X}$. On montrera plus bas, moyennant le calcul de ses composantes dans une carte arbitraire (U, h), que *le champ* $\pounds_X S$ *est différentiable.*

Définition 4. Le champ $\pounds_X S$ s'appelle *dérivée de Lie* du champ de tenseurs S par rapport au champ de vecteurs X.

Il est aisé de voir que *pour chaque champ de vecteurs X l'application $\pounds_X$ est une dérivation de l'algèbre des champs de tenseurs sur la variété $\mathscr{X}$*, i.e. $\pounds_X$ est linéaire, et pour tous champs de tenseurs S et T on a la formule

$$\pounds_X (S \otimes T) = \pounds_X S \otimes T + S \otimes \pounds_X T. \tag{13}$$

En effet, la linéarité est évidente ; pour prouver la formule (13) il suffit de remarquer qu'en vertu de la formule (11)

$$\varphi_t^* (S \otimes T) - S \otimes T = (\varphi_t^* S - S) \otimes \varphi_t^* T + S \otimes (\varphi_t^* T - T)$$

pour tout t. □

Exercice 6. Montrer que *l'opération $\pounds_X$ commute à l'opération de convolution des champs de tenseurs* (par rapport à tout couple d'indices).

Si un champ S est une fonction différentiable f, d'après la formule (12)

$$(\pounds_X f)(p) = \lim_{t \to 0} \frac{f(\varphi_t(p)) - f(p)}{t}$$

pour tout point $p \in \mathscr{X}$. Soit $(U, h) = (U, x^1, \ldots, x^n)$ une carte de $\mathscr{X}$ telle que $p \in U$, et soit $f = f(x^1, \ldots, x^n)$ sur U. Si $x^i = x^i(t)$, $i = 1, \ldots, n$, sont les équations paramétriques de la

courbe intégrale $t \mapsto \varphi_t(p)$ du champ X dans la carte (U, h), alors

$$\lim_{t\to 0} \frac{f(\varphi_t(p)) - f(p)}{t} = \lim_{t\to 0} \frac{f(x^1(t), \ldots, x^n(t)) - f(x^1(0), \ldots, x^n(0))}{t} =$$

$$= \frac{\mathrm{d}f(x^1(t), \ldots, x^n(t))}{\mathrm{d}t}\Big|_{t=0} = \left(\frac{\partial f}{\partial x^i}\right)_p \frac{\mathrm{d}x^i(t)}{\mathrm{d}t}\Big|_{t=0} =$$

$$= \left(\frac{\partial f}{\partial x^i}\right)_p X_p^i = (Xf)(p),$$

où X^i, $i = 1, \ldots, n$, sont les composantes du champ de vecteurs X dans la carte (U, h).

Ceci prouve que $\pounds_X f = Xf$ sur U, donc, puisque la carte (U, h) est arbitraire, que

$$\pounds_X f = Xf \quad \text{sur} \quad \mathscr{X}. \tag{14}$$

Donc, *l'opération $\pounds_X$ est une généralisation de l'opération X avec des fonctions sur des champs de tenseurs arbitraires.*

En vertu de la formule (13) il s'ensuit en particulier que

$$\pounds_X (fS) = Xf \cdot S + f\pounds_X S \tag{15}$$

pour toute fonction f et tout champ de tenseurs S.

Supposons maintenant que le champ S est un champ de covecteurs α, donc, en vertu de la formule (10),

$$(\varphi_t^* \alpha)_p = (\mathrm{d}\varphi)_p^* \, \alpha_{\varphi(p)}.$$

Par conséquent, la valeur $(\varphi_t^* \alpha)_p A$ du covecteur $\varphi_t^* \alpha$ sur un vecteur $A \in \mathrm{T}_p \mathscr{X}$ s'exprime par la formule

$$(\varphi_t^* \alpha)_p A = \alpha_{\varphi(p)} ((\mathrm{d}\varphi_t)_p A).$$

Donc, si dans la carte (U, h)

$$A = a^i \left(\frac{\partial}{\partial x^i}\right)_p \quad \text{et} \quad \alpha = \alpha_j \, \mathrm{d}x^j,$$

et l'application φ_t est définie par les fonctions

$$y^k = \varphi_t^k (x^1, \ldots, x^n),$$

alors

$$(\varphi_t^* \alpha)_p A = [\alpha_j (\varphi_t(p)) (\mathrm{d}x^j)_{\varphi_t(p)}] \left(a^i \left(\frac{\partial \varphi_t^k}{\partial x^i}\right)_p \left(\frac{\partial}{\partial x^k}\right)_{\varphi_t(p)}\right) =$$

$$= \alpha_j (\varphi_t(p)) \, a^i \left(\frac{\partial \varphi_t^j}{\partial x^i}\right)_p$$

et en particulier

$$(\varphi_t^* \alpha)_p \left(\frac{\partial}{\partial x^i}\right)_p = \alpha_j (\varphi_t(p)) \left(\frac{\partial \varphi_t^j}{\partial x^i}\right)_p.$$

Puisque pour tout covecteur $\xi \in T_p^* \mathscr{X}$

$$\xi = \xi \left(\frac{\partial}{\partial x^i} \right)_p (dx^i)_p,$$

ceci prouve que

$$(\varphi_t^* \alpha)_p = (\alpha_j \circ \varphi_t)(p) \left(\frac{\partial \varphi_t^j}{\partial x^i} \right)_p (dx^i)_p.$$

Donc,

$$(\varphi_t^* \alpha)_p - \alpha_p = \left[(\alpha_j \circ \varphi_t)(p) \left(\frac{\partial \varphi_t^j}{\partial x^i} \right)_p - \alpha_i(p) \right] (dx^i)_p$$

et, par suite,

$$(\pounds_X \alpha)_p = \beta_i(p) (dx^i)_p,$$

où

$$\beta_i(p) = \lim_{t \to 0} \frac{(\alpha_j \circ \varphi_t)(p) \left(\frac{\partial \varphi_t^j}{\partial x^i} \right)_p - \alpha_i(p)}{t}.$$

Supposons en particulier que $\alpha = dx^k$, i.e. $\alpha_j = \delta_j^k$. Alors $(\alpha_j \circ \varphi_t)(p) = \delta_j^k$ pour tout t et, par suite,

$$\beta_i(p) = \lim_{t \to 0} \frac{\left(\frac{\partial \varphi_t^k}{\partial x^i} \right)_p - \delta_i^k}{t} = \frac{\partial}{\partial t} \left(\frac{\partial \varphi_t^k}{\partial x^i} \right)_p \Big|_{t=0} = \left(\frac{\partial}{\partial x^i} \left(\frac{\partial \varphi_t^k}{\partial t} \Big|_{t=0} \right) \right)_p.$$

(Remarquons que $\delta_i^k = \left(\frac{\partial \varphi_0^k}{\partial x^i} \right)_p$, car $\varphi_0^k(x^1, \ldots, x^n) = x^k$.) Mais si comme plus haut $x^i = x^i(t)$ sont les équations paramétriques d'une courbe $t \mapsto \varphi_t(p)$ dans la carte (U, h), alors par définition

$$x^i(t) = \varphi_t^i(x_0^1, \ldots, x_0^n), \quad i = 1, \ldots, n,$$

où $x_0^1, \ldots, x_0^n$ sont les coordonnées du point p, et donc

$$\frac{\partial \varphi_t^k}{\partial t} \Big|_{t=0} = \frac{dx^k(t)}{dt} \Big|_{t=0} = X^k(p), \quad k = 1, \ldots, n,$$

où X^k, $k = 1, \ldots, n$, sont comme plus haut les composantes du champ de vecteurs X dans la carte (U, h). Donc, dans le cas envisagé

$$\beta_i(p) = \left(\frac{\partial X^k}{\partial x^i} \right)_p.$$

Ceci prouve que

$$\pounds_X dx^k = \frac{\partial X^k}{\partial x^i} dx^i \text{ sur } U$$

et, par suite (cf. formule (15)),

$$\pounds_X(\alpha_k\,\mathrm{d}x^k) = X\alpha_k\,\mathrm{d}x^k + \alpha_k\,\frac{\partial X^k}{\partial x^i}\,\mathrm{d}x^i,$$

i.e.

$$\pounds_X\alpha = \left(X\alpha_i + \alpha_k\,\frac{\partial X^k}{\partial x^i}\right)\mathrm{d}x^i \text{ sur } U, \tag{16}$$

ou sous la forme plus symétrique

$$\pounds_X\alpha = \left(X^k\,\frac{\partial\alpha_i}{\partial x^k} + \alpha_k\,\frac{\partial X^k}{\partial x^i}\right)\mathrm{d}x^i \text{ sur } U.$$

On voit en particulier — conformément à la proposition générale ci-dessus — que le champ de covecteurs $\pounds_X\alpha$ est différentiable.

La dérivée de Lie $\pounds_X Y$ du champ de vecteurs Y par rapport au champ de vecteurs X peut être obtenue par un calcul identique. Mais dans le cas d'un champ de covecteurs α arbitraire, il est plus simple d'envisager le champ de tenseurs $\alpha \otimes Y$ sur la variété $\mathscr{X}$ et de se servir du fait, *primo*, qu'en vertu de la formule générale (11)

$$\pounds_X(\alpha\otimes Y) = \pounds_X\alpha\otimes Y + \alpha\otimes\pounds_X Y$$

et, *secundo*, que l'opération $\pounds_X$ commute à l'opération de convolution des champs de tenseurs par rapport à tout couple d'indices (cf. exercice 5 ci-dessus). Puisque le convolé du champ $\alpha \otimes Y$ est visiblement la fonction différentiable

$$\alpha(Y):\ p \mapsto \alpha_p(Y_p),$$

on en déduit que

$$\pounds_X[\alpha(Y)] = (\pounds_X\alpha)(Y) + \alpha(\pounds_X Y). \tag{17}$$

Dans une carte arbitraire (U, h) la fonction $\alpha(Y)$ s'exprime de toute évidence à l'aide de la formule

$$\alpha(Y) = \alpha_i Y^i \quad \text{sur} \quad U,$$

où α_i et Y^i, $i = 1, \ldots, n$, sont les composantes respectives des champs α et Y et donc, en vertu des formules (14) et (16), la formule (17) devient

$$X[\alpha_i Y^i] = \left(X\alpha_i + \alpha_k\,\frac{\partial X^k}{\partial x^i}\right)Y^i + \alpha_i Z^i \text{ sur } U,$$

où Z^i, $i = 1, \ldots, n$, sont les composantes du champ de vecteurs $\pounds_X Y$ dans la carte (U, h).

En particulier, pour $\alpha_i = \delta_i^j$ on en déduit que

$$XY^j = \frac{\partial X^j}{\partial x^i}\,Y^i + Z^j,$$

i.e.

$$Z^j = XY^j - \frac{\partial X^j}{\partial x^i} Y^i = X^i \frac{\partial Y^j}{\partial x^i} - \frac{\partial X^j}{\partial x^i} Y^i.$$

En comparant ce résultat avec la formule (22) de la leçon précédente, on découvre aussitôt que $Z^j = [X, Y]^j$ et, par suite,

$$\text{(18)} \qquad \pounds_X Y = [X, Y].$$

Dans une algèbre de Lie $\mathfrak{g}$, l'application

$$x \longrightarrow [a, x], \quad a, x \in \mathfrak{g},$$

est désignée par le symbole ad a. En se servant de cette notation, on peut écrire (18) sous la forme suivante:

$$\pounds_X = \operatorname{ad} X \quad \text{sur } \alpha\mathscr{X}.$$

Nous avons ainsi appris à calculer $\pounds_X$ sur des fonctions, des champs de covecteurs et des champs de vecteurs. Etant donné qu'un champ de tenseurs arbitraire S s'exprime dans (U, h) par la formule

$$S = S^{j_1 \ldots j_b}_{i_1 \ldots i_a} \, dx^{i_1} \otimes \ldots \otimes dx^{i_a} \otimes \frac{\partial}{\partial x^{j_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{j_b}},$$

on peut, en se servant de la formule (11) (et de son cas particulier (15)), calculer toutes les composantes $(\pounds_X S)^{j_1 \ldots j_b}_{i_1 \ldots i_a}$ du champ $\pounds_X S$ dans la carte (U, h). Sans détailler la formule correspondante, on peut affirmer *a priori* qu'elle exprime les composantes $(\pounds_X S)^{j_1 \ldots j_b}_{i_1 \ldots i_a}$ sous forme d'une somme de produits des composantes de S et de X par leurs dérivées. Ces composantes sont donc des fonctions différentiables et, par suite, le champ $\pounds_X S$ est différentiable.

Exercice 7. Montrer que

$$(\pounds_X S)^{j_1 \ldots j_b}_{i_1 \ldots i_a} = \frac{\partial S^{j_1 \ldots j_b}_{i_1 \ldots i_a}}{\partial x^k} X^k +$$

$$+ S^{j_1 j_2 \ldots j_b}_{k i_2 \ldots i_a} \frac{\partial X^k}{\partial x^{i_1}} + S^{j_1 j_2 \ldots j_b}_{i_1 k \ldots i_a} \frac{\partial X^k}{\partial x^{i_2}} + \ldots + S^{j_1 \ldots j_{b-1} j_b}_{i_1 \ldots i_{a-1} k} \frac{\partial X^k}{\partial x^{i_a}} -$$

$$- S^{k j_2 \ldots j_b}_{i_1 i_2 \ldots i_a} \frac{\partial X^{j_1}}{\partial x^k} - S^{j_1 k \ldots j_b}_{i_1 i_2 \ldots i_a} \frac{\partial X^{j_2}}{\partial x^k} - \ldots - S^{j_1 \ldots j_{b-1} k}_{i_1 \ldots i_{a-1} i_a} \frac{\partial X^{j_b}}{\partial x^k}.$$

LEÇON 18

Formes différentielles linéaires. — Formes différentielles de degré quelconque. — Formes différentielles en tant que fonctionnelles de champs de vecteurs. — Produit intérieur d'un champ de vecteurs par une forme différentielle. — Transfert d'une forme différentielle par une application différentiable.

Etudions maintenant les champs de covecteurs sur $\mathscr{X}$. Un tel champ α s'écrit dans une carte $(U, x^1, \ldots, x^n)$ sous la forme

$$\alpha = \alpha_i \, dx^i \quad \text{sur } U$$

(cf. formule (13) de la leçon 16), i.e. se représente par une forme linéaire des différentielles $dx^1, \ldots, dx^n$ des coordonnées locales. Pour cette raison, les champs de covecteurs sont généralement appelés *formes différentielles linéaires* (et les fonctions α_i, $i = 1, \ldots, n$, *coefficients* de ces formes dans la carte (U, h)). L'espace vectoriel $\mathsf{T}_1\mathscr{X}$ des formes différentielles linéaires se désigne aussi par le symbole $\Omega^1\mathscr{X}$.

Pour toute forme différentielle linéaire α et tout champ de vecteurs X la formule

$$\alpha(X)(p) = \alpha_p(X_p), \quad p \in \mathscr{X},$$

définit une fonction $\alpha(X)$ sur $\mathscr{X}$. Cette fonction est désignée aussi par les symboles $\langle \alpha, X \rangle$, $i_X\alpha$ et $X \lrcorner \alpha$, et s'appelle *produit intérieur* de la forme α et du champ X. Dans chaque carte $(U, h) = (U, x^1, \ldots, x^n)$ la fonction $\alpha(X)$ s'exprime par

$$\alpha(X) = \alpha_i X^i, \tag{1}$$

où X^i sont les composantes du champ X, et α_i les coefficients de la forme α dans la carte (U, h), d'où il résulte immédiatement que *la fonction $\alpha(X)$ est différentiable sur $\mathscr{X}$*. [A la leçon 17 nous avons déjà considéré cette fonction lors du calcul de la dérivée de Lie d'un champ de vecteurs.]

Donc, pour toute forme différentielle $\alpha \in \Omega^1\mathscr{X}$, la formule

$$X \mapsto \alpha(X)$$

définit une application

$$\alpha : \mathfrak{a}\mathscr{X} \to \mathsf{F}\mathscr{X} \tag{2}$$

qui est de toute évidence un morphisme de $\mathsf{F}\mathscr{X}$-*modules*, i.e. vérifie la relation

$$\alpha(fX) = f\alpha(X)$$

pour toute fonction $f \in \mathsf{F}\mathscr{X}$ et tout champ $X \in \mathfrak{a}\mathscr{X}$.

On dira que le morphisme (2) est *engendré* par la forme différentielle α.

Proposition 1. *Pour toute variété séparée différentiable $\mathscr{X}$ la correspondance*

(3) $$\textit{forme } \alpha \Rightarrow \textit{morphisme } (2)$$

est un isomorphisme de l'espace vectoriel $\Omega^1\mathscr{X}$ sur l'espace vectoriel $\mathrm{Hom}_{\mathsf{F}\mathscr{X}}(\mathfrak{a}\mathscr{X}, \mathsf{F}\mathscr{X})$ *des morphismes* (2).

La démonstration de cette proposition passe par celle de quelques lemmes simples sur les morphismes (2). On admettra dans ces lemmes que la variété $\mathscr{X}$ est séparée.

On dira que des champs de vecteurs X et Y sont *confondus à proximité d'un point* $p \in \mathscr{X}$ s'ils le sont dans un voisinage de ce point. (Cf. dans la leçon 16 une définition analogue pour les fonctions.)

Lemme 1 (propriété locale des morphismes $\mathfrak{a}\mathscr{X} \to \mathsf{F}\mathscr{X}$). *Si des champs de vecteurs X et Y sont confondus à proximité d'un point $p \in \mathscr{X}$, pour tout morphisme* (2) *les fonctions $\alpha(X)$ et $\alpha(Y)$ le sont aussi.*

D é m o n s t r a t i o n. (Comparer avec la démonstration du corollaire 2 au lemme 1 de la leçon 16.) Supposons que $X = Y$ dans un voisinage U du point p et soit φ une fonction différentiable $\mathscr{X} \to \mathbb{R}$ égale à 0 à l'extérieur de U et à 1 dans un voisinage $W \subset U$ de p (cf. corollaire 1 au lemme 1 de la leçon 16). Le champ $\varphi\cdot(X-Y)$ est alors nul sur $\mathscr{X}$, i.e. est le zéro de l'espace vectoriel $\mathfrak{a}\mathscr{X}$. Donc, en vertu de la linéarité du morphisme α,

$$\alpha[\varphi\cdot(X-Y)] = 0$$

et, par conséquent, $\varphi\cdot\alpha(X-Y) = 0$. Puisque $\varphi = 1$ sur W, ceci prouve que $\alpha(X-Y) = 0$ sur W, donc, $\alpha(X) = \alpha(Y)$ sur W. □

Lemme 2. *Supposons que U est ouvert dans $\mathscr{X}$ et soit $p_0 \in U$. Pour tout champ de vecteurs X' sur U, il existe alors un champ de vecteurs X sur $\mathscr{X}$ tel que $X' = X$ à proximité p_0.*

D é m o n s t r a t i o n. (Comparer avec celle du lemme 1 de la leçon 16.) D'après la proposition 2 de la leçon 14, il existe dans $\mathscr{X}$ des ensembles ouverts V et W tels que

$$p_0 \in W, \quad \overline{W} \subset V, \quad \overline{V} \subset U,$$

et pour le couple (V, W) existe la fonction d'Urysohn φ. Considérons un point quelconque $p \in \mathscr{X}$ et posons

$$X_p = \begin{cases} \varphi(p)\, X'_p & \text{si } p \in U, \\ 0 & \text{si } p \notin U. \end{cases}$$

Il est clair que le champ $X: p \mapsto X_p$ est différentiable sur $\mathscr{X}$ et coïncide sur W avec le champ X'. □

Corollaire 1. *Pour tout point p_0 de la variété $\mathscr{X}$ et tout vecteur $A \in \mathrm{T}_{p_0}\mathscr{X}$, il existe un champ de vecteurs X sur $\mathscr{X}$ tel que*

$$X_{p_0} = A.$$

Démonstration. Soit $(U, h) = (U, x^1, \ldots, x^n)$ une carte telle que $p_0 \in U$ et soit

$$A = a^i \left(\frac{\partial}{\partial x^i}\right)_{p_0}.$$

Définissons un champ de vecteurs X' sur U en posant

$$X'_p = a^i \left(\frac{\partial}{\partial x^i}\right)_p$$

pour tout point $p \in U$. D'après le lemme 2, il existe un champ X sur $\mathscr{X}$ tel que $X' = X$ à proximité de p_0. En particulier, $X_{p_0} = X'_{p_0} = A$. □

Corollaire 2. *Pour tout ensemble ouvert $U \subset \mathscr{X}$, tout morphisme $\alpha: \mathfrak{a}\mathscr{X} \to \mathsf{F}\mathscr{X}$ induit un seul morphisme*

$$\alpha_U : \mathfrak{a}U \to \mathsf{F}U, \tag{4}$$

qui est justiciable du diagramme commutatif

$$\begin{array}{ccc} \mathfrak{a}\mathscr{X} & \to & \mathfrak{a}U \\ {\scriptstyle\alpha}\downarrow & & \downarrow{\scriptstyle\alpha_U} \\ \mathsf{F}\mathscr{X} & \to & \mathsf{F}U, \end{array} \tag{5}$$

dont les flèches horizontales figurent des applications de restriction, i.e. tel que

$$\alpha(X)|_U = \alpha_U(X|_U)$$

pour tout champ de vecteurs X sur $\mathscr{X}$. (Comparer avec la proposition 1 de la leçon 16.)

Démonstration. Supposons que le morphisme α_U existe et soit X un champ de vecteurs arbitraire sur U. En vertu du lemme 2, pour tout point $p \in U$ il existe un champ de vecteurs X' sur $\mathscr{X}$ confondu avec X à proximité de p. Ceci étant, si $X' = X$ au voisinage W de p, alors

$$\alpha_U(X)|_W = \alpha(X')|_W$$

et, en particulier,

$$\alpha_U(X)(p) = \alpha(X')(p). \tag{6}$$

En outre, d'après la propriété locale du morphisme α, pour deux champs quelconques X' et X'' sur $\mathscr{X}$ confondus à proximité de p avec X, on aura l'égalité

$$\alpha(X')(p) = \alpha(X'')(p) \tag{6'}$$

qui prouve que le second membre de la formule (6) est indépendant du choix de X. Ceci prouve l'unicité du morphisme α_U.

Pour démontrer l'existence du morphisme α_U, définissons pour tout champ de vecteurs X sur U la fonction $\alpha_U(X)$ à l'aide de la formule (6). Cette définition est cohérente en vertu de la formule (6'). De plus, si $X' = X$ au voisinage W de p, alors $\alpha_U(X) = \alpha(X')$ sur W, d'où il résulte que la fonction $\alpha_U(X)$ est différentiable sur U, donc, la formule $\alpha_U : X \mapsto \alpha_U(X)$ définit une application

$$\alpha_U : \mathfrak{a}U \to \mathsf{F}U.$$

Si maintenant $X \in \mathfrak{a}U$ et $f \in \mathsf{F}U$, tandis que $X_1 \in \mathfrak{a}\mathscr{X}$ et $f_1 \in \mathsf{F}\mathscr{X}$, et de plus $X = X_1$ et $f = f_1$ à proximité de p, alors $fX = f_1X_1$ à proximité de p, donc

$$\alpha_U(fX)(p) = \alpha(f_1X_1)(p) = (f_1\alpha(X))(p) =$$
$$= f_1(p)\cdot\alpha(X_1)(p) = f(p)\cdot\alpha_U(X)(p) = [f\alpha_U(X)](p).$$

Donc, $\alpha_U(fX) = f\alpha_U(X)$ et, par suite, α_U est un morphisme de modules.

Enfin, tout champ X sur $\mathscr{X}$ étant susceptible de servir de champ X' pour le champ $X|_U$ (par rapport à un point quelconque $p \in U$), on a

$$\alpha_U(X|_U)(p) = \alpha(X)(p)$$

en tout point $p \in U$. Donc, le diagramme (5) est commutatif. □

Lemme 3. *Dans toute carte* $(U, h) = (U, x^1, \ldots, x^n)$, *tout morphisme* (4) *agit d'après la formule*

$$\alpha_U(X) = \alpha_i X^i, \tag{7}$$

où α_i, $i = 1, \ldots, n$, *sont des fonctions différentiables sur* U, *et* X^i, $i = 1, \ldots, n$, *les composantes du champ de vecteurs* X *dans la carte* (U, h).

Démonstration. Puisque $X = X^i \frac{\partial}{\partial x^i}$ sur U, on a

$$\alpha_U(X) = \alpha_U\left(\frac{\partial}{\partial x^i}\right)X^i = \alpha_i X^i,$$

où

$$\alpha_i = \alpha\left(\frac{\partial}{\partial x^i}\right). \quad \square$$

Corollaire 1. *Le morphisme* (4) *est engendré sur* U *par la forme différentielle*

$$\alpha_U = \alpha_i\, dx^i. \quad \square \tag{8}$$

Désormais nous sommes en mesure de prouver la proposition 1.

Démonstration de la proposition 1. Pour ne pas nous embrouiller dans les identifications, dans cette démons-

tration nous désignerons par $\hat{\alpha}$ le morphisme (2) engendré par la forme α. Il est clair que l'application (3) est linéaire. Supposons que $\hat{\alpha} = 0$, i.e. $\alpha_p(X_p) = 0$ pour tout champ de vecteurs X sur $\mathscr{X}$ et tout point $p \in \mathscr{X}$. En vertu du corollaire 1 du lemme 2, pour tout point $p_0 \in \mathscr{X}$ et tout vecteur $A \in \mathrm{T}_{p_0}\mathscr{X}$ il existe un champ $X \in \alpha\mathscr{X}$ tel que $X_{p_0} = A$. Donc, $\alpha_{p_0}(A) = 0$, i.e. $\alpha_{p_0} = 0$ sur $\mathrm{T}_{p_0}\mathscr{X}$. Par conséquent, $\alpha = 0$.

Ceci prouve que l'application (3) est un monomorphisme.

Soit un morphisme arbitraire $\beta: \alpha\mathscr{X} \to \mathsf{F}\mathscr{X}$. En vertu du corollaire 1 du lemme 3, pour toute carte $(U, h) = (U, x^1, \ldots, x^n)$ le morphisme $\beta_U: \alpha U \to \mathsf{F}U$ est engendré par la forme (8), où $\alpha_i = \beta\left(\frac{\partial}{\partial x^i}\right)$. Si $(U', h') = (U', x^{1'}, \ldots, x^{n'})$ est une autre carte et si le morphisme $\beta_{U'}$ sur U' est engendré par la forme $\alpha_{U'} = \alpha_{i'}\, dx^{i'}$, d'après la propriété locale du morphisme β, on aura sur l'intersection $U \cap U'$

$$\alpha_i\, dx^i = \alpha_{i'}\, dx^{i'}.$$

Ceci montre qu'en posant

$$\alpha_p = (\alpha_U)_p, \quad p \in U,$$

on définit de façon unique un champ de covecteurs α sur $\mathscr{X}$ tel que $\alpha|_U = \alpha_U$ pour tout voisinage de coordonnées U et donc, *primo*, α est différentiable, *secundo*, il engendre le morphisme β (i.e. est tel que $\hat{\alpha} = \beta$).

L'application (3) est donc un épimorphisme. □

Dans la suite on identifiera en principe les formes différentielles linéaires (les champs de covecteurs) sur $\mathscr{X}$ et les morphismes (2) qu'elles engendrent.

Remarque 1. Attirons l'attention sur le fait que *l'application* (3) *est un morphisme de* $\mathsf{F}\mathscr{X}$*-modules*, i.e., dans les notations introduites pour démontrer la proposition 1, pour toute forme $\alpha \in \Omega^1\mathscr{X}$ et toute fonction $f \in \mathsf{F}\mathscr{X}$, on a

$$\widehat{f\alpha} = f\hat{\alpha},$$

où le morphisme $f\hat{\alpha}$ est défini, selon l'usage en algèbre, par la formule $(f\hat{\alpha})(X) = f\hat{\alpha}(X)$.

* * *

Les champs de tenseurs ω associant à tout point $p \in \mathscr{X}$ un *tenseur antisymétrique* ω_p, i.e. (cf. leçon II.8) un tenseur de type $(r, 0)$ dont les composantes changent de signe par toute transposition des indices, sont d'une très grande importance. (Le nombre r s'appelle

degré du champ ω.) Pour tous champs θ et ω de tenseurs antisymétriques, la formule

$$(\theta \wedge \omega)_p = \theta_p \wedge \omega_p,$$

où $\theta_p \wedge \omega_p$ est le produit extérieur des tenseurs θ_p et ω_p (cf. leçon II.8) définit un champ $\theta \wedge \omega$ de tenseurs antisymétriques dont le degré est égal à la somme des degrés de θ et ω. *Si les champs* θ *et* ω *sont différentiables, il en est de même de* $\theta \wedge \omega$, puisque les composantes du produit extérieur s'expriment algébriquement en fonction des composantes des facteurs.

Les propriétés algébriques du produit extérieur de tenseurs antisymétriques (par exemple, l'associativité et l'anticommutativité; cf. leçon II.8) sont toutes valables pour leurs champs. Donc, dans le produit extérieur d'un nombre quelconque de champs de tenseurs antisymétriques, on peut omettre d'écrire les parenthèses et, pour tous champs θ et ω de tenseurs antisymétriques, on a la formule

$$\omega \wedge \theta = (-1)^{rs}\, \theta \wedge \omega,$$

où r et s sont les degrés respectifs des champs ω et θ.

Pour $r = 0$, le champ ω est une fonction f, et le produit extérieur $\omega \wedge \theta$, le produit ordinaire $f\theta$ de la fonction f par le champ θ.

Les expressions classiques des tenseurs antisymétriques de type $(r, 0)$ en fonction des produits extérieurs des covecteurs de la base duale (cf. formule (8) et proposition (8) de la leçon II.8) montrent que chaque champ de tenseurs antisymétriques sur un voisinage de coordonnées U s'exprime par la formule

$$(9) \qquad \omega = \frac{1}{r!}\, \omega_{i_1 \ldots i_r}\, dx^{i_1} \wedge \ldots \wedge dx^{i_r} = \\ = \sum_{1 \leqslant i_1 < \ldots < i_n \leqslant n} \cdots \sum \omega_{i_1 \ldots i_n}\, dx^{i_1} \wedge \ldots \wedge dx^{i_r},$$

où $\omega_{i_1 \ldots i_r}$ sont les composantes du champ ω dans la carte $(U, x^1, \ldots, x^n)$ (qui sont des fonctions différentiables sur U). Ceci explique que les champs de tenseurs antisymétriques soient aussi appelés *formes différentielles sur la variété* $\mathcal{X}$.

Pour $r = 1$, on obtient les formes différentielles linéaires étudiées plus haut, pour $r = 0$, des fonctions différentiables sur $\mathcal{X}$.

Toutes les formes différentielles de degré $r \geqslant 0$ engendrent un sous-espace vectoriel $\Omega^r\mathcal{X}$ de l'espace $\mathsf{T}_r\mathcal{X}$ des champs de tenseurs de type $(r, 0)$. Donc, $\Omega^0\mathcal{X} = \mathsf{F}\mathcal{X}$ et $\Omega^1\mathcal{X} = \mathsf{T}_1\mathcal{X}$ (alors que pour $r > 1$, l'inclusion $\Omega^r\mathcal{X} \subset \mathsf{T}_r\mathcal{X}$ est visiblement stricte). On s'est déjà servi plus haut du symbole $\Omega^1\mathcal{X}$.

Remarquons que

$$\Omega^r\mathcal{X} = 0$$

pour tout $r > n$.

* * *

En traitant les tenseurs de type $(r, 0)$ comme des fonctionnelles multilinéaires de vecteurs, on peut associer à toute forme différentielle ω de degré r et à tous champs de vecteurs $X_1, \ldots, X_r$ une fonction $\omega(X_1, \ldots, X_r)$ sur $\mathscr{X}$ dont la valeur en un point $p \in \mathscr{X}$ est définie par la formule

$$\omega(X_1, \ldots, X_r)(p) = \omega_p((X_1)_p, \ldots, (X_r)_p).$$

Si dans la carte $(U, x^1, \ldots, x^n)$

$$\omega = \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \ldots \sum \omega_{i_1 \ldots i_r} \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_r}$$

et

$$X_1 = X_1^{i_1} \frac{\partial}{\partial x^{i_1}}, \ \ldots, \ X_r = X_r^{i_r} \frac{\partial}{\partial x^{i_r}},$$

alors

$$\omega(X_1, \ldots, X_r) = \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \ldots \sum \omega_{i_1 \ldots i_r} \begin{vmatrix} X_1^{i_1} & \ldots & X_1^{i_r} \\ \ldots & \ldots & \ldots \\ X_r^{i_1} & \ldots & X_r^{i_r} \end{vmatrix} \text{ sur } U$$

(cf. formule (5) de la leçon II.8). Donc, *la fonction* $\omega(X_1, \ldots, X_r)$ *est différentiable.*

L'application

$$(10) \quad \omega : \underbrace{\mathfrak{a}\mathscr{X} \times \ldots \times \mathfrak{a}\mathscr{X}}_{r \text{ fois}} \to \mathsf{F}\mathscr{X}, \ (X_1, \ldots, X_r) \mapsto \omega(X_1, \ldots, X_r),$$

est de toute évidence $\mathsf{F}\mathscr{X}$*-multilinéaire*, i.e. est un morphisme de $\mathsf{F}\mathscr{X}$-modules par rapport à chaque argument. De plus, elle est *antisymétrique*, i.e. change de signe par toute transposition des arguments.

De même que pour $r = 1$ (cf. proposition 1 et remarque 1), *si une variété* $\mathscr{X}$ *est séparée, pour tout* $r \geqslant 1$ *la correspondance*

(11) *forme de degré* $r \Rightarrow$ *application* (10)

définit un isomorphisme du $\mathsf{F}\mathscr{X}$*-module* $\Omega^r\mathscr{X}$ *sur le* $\mathsf{F}\mathscr{X}$*-module de toutes les applications* $\mathsf{F}\mathscr{X}$*-multilinéaires antisymétriques* (10). La démonstration qui reprend *ad litteram* celle de la proposition 1 est laissée au soin du lecteur (qui est vivement conseillé de l'effectuer en détail !). Dans la suite, on identifiera d'une façon générale les formes différentielles avec les applications correspondantes (10).

Remarque 2. De façon analogue, des champs de tenseurs quelconques sur $\mathscr{X}$ de type $(r, 0)$, $r > 0$, s'identifient à des applications

$\mathsf{F}\mathscr{X}$-multilinéaires (pas forcément antisymétriques)

$$\underbrace{\mathfrak{a}\mathscr{X} \times \ldots \times \mathfrak{a}\mathscr{X}}_{r \text{ fois}} \to \mathsf{F}\mathscr{X},$$

et les champs de tenseurs de type (r, s), $r \geqslant 0$, $s \geqslant 0$, à des applications $\mathsf{F}\mathscr{X}$-multilinéaires de la forme

$$(12) \qquad S : \underbrace{\mathfrak{a}\mathscr{X} \times \ldots \times \mathfrak{a}\mathscr{X}}_{r \text{ fois}} \times \underbrace{\Omega^1\mathscr{X} \times \ldots \times \Omega^1\mathscr{X}}_{s \text{ fois}} \to \mathsf{F}\mathscr{X}.$$

En particulier,

$$\mathfrak{a}\mathscr{X} = \mathrm{Hom}_{\mathsf{F}\mathscr{X}}(\Omega^1\mathscr{X}, \mathsf{F}\mathscr{X})$$

(au champ X est associée une application $\mathsf{F}\mathscr{X}$-linéaire $i_X : \alpha \mapsto X \lrcorner\, \alpha$). Donc, pour $s = 1$, chaque champ (12) peut être traité comme une application $\mathsf{F}\mathscr{X}$-multilinéaire

$$S : \underbrace{\mathfrak{a}\mathscr{X} \times \ldots \times \mathfrak{a}\mathscr{X}}_{r \text{ fois}} \to \mathfrak{a}\mathscr{X},$$

associant aux champs de vecteurs $X_1, \ldots, X_r$ un champ de vecteurs $S(X_1, \ldots, X_r)$ tel que

$$S(X_1, \ldots, X_v) \lrcorner\, \alpha = S(X_1, \ldots, X_r, \alpha)$$

pour toute forme différentielle $\alpha \in \Omega^1\mathscr{X}$.

Exercice 1. Donner une interprétation identique de l'application (12) pour $s > 1$.

* * *

En vertu de l'identification (11), chaque champ de vecteurs X permet d'associer à une forme quelconque ω de degré $r > 0$ la forme $i_X\omega = X \lrcorner\, \omega$ de degré $r - 1$, dont la valeur sur les champs de vecteurs $X_1, \ldots, X_{r-1}$ est définie par la formule

$$(X \lrcorner\, \omega)(X_1, \ldots, X_{r-1}) = \omega(X, X_1, \ldots, X_{r-1}).$$

La forme $X \lrcorner\, \omega$ s'appelle *produit intérieur* du champ X par la forme ω. [Pour $r = 1$, nous l'avons déjà étudiée au début de cette leçon.]

Pour $r = 0$ (i.e. dans le cas où la forme ω est une fonction) on admettra par définition que $X \lrcorner\, \omega = 0$ pour tout champ X.

On appelle *battage* σ *de type* (r, s) une permutation de degré $n = r + s$ qui préserve séparément l'ordre respectif des nombres $1, \ldots, r$ et des nombres $r + 1, \ldots, r + s$, i.e. telle que

$$\sigma(1) < \ldots < \sigma(r) \quad \text{et} \quad \sigma(r + 1) < \ldots < \sigma(r + s).$$

On démontre (faites-le!) que, lorsque les formes différentielles sont traitées comme des applications (10), le produit extérieur $\theta \wedge \omega$

d'une forme θ de degré r par une forme ω de degré s est défini par la formule

$$(13)\quad (\theta \wedge \omega)(X_1, \ldots, X_{r+s}) = \\ = \sum_\sigma \varepsilon_\sigma \theta(X_{\sigma(1)}, \ldots, X_{\sigma(r)})\, \omega(X_{\sigma(r+1)}, \ldots, X_{\sigma(r+s)}),$$

où la sommation est étendue à tous les battages σ de type (r, s).

Il s'ensuit que *le produit intérieur* $X \lrcorner (\theta \wedge \omega)$ *du champ de vecteurs* X *par la forme* $\theta \wedge \omega$ *est justiciable de la formule*

$$(14)\qquad X \lrcorner (\theta \wedge \omega) = (X \lrcorner \theta) \wedge \omega + (-1)^r \theta \wedge (X \lrcorner \omega),$$

où r *est le degré de la forme* θ. En effet, en vertu de la formule (13) pour tous champs de vecteurs $X_1, \ldots, X_{r+s-1}$

$$X \lrcorner (\theta \wedge \omega)(X_1, \ldots, X_{r+s-1}) = (\theta \wedge \omega)(X, X_1, \ldots, X_{r+s-1}) = \\ = \sum_\sigma \varepsilon_\sigma \theta(Y_{\sigma(1)}, \ldots, Y_{\sigma(r)})\, \omega(Y_{\sigma(r+1)}, \ldots, Y_{\sigma(r+s)}),$$

$$((X \lrcorner \theta) \wedge \omega)(X_1, \ldots, X_{r+s-1}) = \\ = \sum_\sigma \varepsilon_\sigma \theta(Y_1, Y_{\sigma(2)}, \ldots, Y_{\sigma(r)})\, \omega(Y_{\sigma(r+1)}, \ldots, Y_{\sigma(r+s)}),$$

où $Y_1 = X$, $Y_2 = X_1, \ldots, Y_{r+s} = X_{r+s-1}$ et où dans la première somme σ parcourt tous les battages de type (r, s) et dans la deuxième seulement ceux pour lesquels $\sigma(1) = 1$. Puisque pour tout battage σ de type (r, s) soit $\sigma(1) = 1$, soit $\sigma(r+1) = 1$, on en déduit que

$$(15)\quad (X \lrcorner (\theta \wedge \omega) - (X \lrcorner \theta) \wedge \omega)(X_1, \ldots, X_{r+s-1}) = \\ = \sum_\sigma \varepsilon_\sigma \theta(Y_{\sigma(1)}, \ldots, Y_{\sigma(r)})\, \omega(Y_1, Y_{\sigma(r+2)}, \ldots, Y_{\sigma(r+s)}),$$

où la sommation est étendue à tous les battages σ tels que $\sigma(r+1) = 1$.

Mais chaque battage σ de type (r, s) tel que $\sigma(r+1) = 1$ définit à l'aide de la formule

$$\tau(a) = \begin{cases} \sigma(a) - 1 & \text{si } a = 1, \ldots, r, \\ \sigma(a+1) - 1 & \text{si } a = r+1, \ldots, r+s-1, \end{cases}$$

un battage τ de type $(r, s-1)$ pour lequel

$$Y_{\sigma(1)} = X_{\tau(1)}, \ldots, Y_{\sigma(r)} = X_{\tau(r)}, \\ Y_{\sigma(r+2)} = X_{\tau(r+1)}, \ldots, Y_{\sigma(r+s)} = X_{\tau(r+s-1)}$$

et

$$\varepsilon_\sigma = (-1)^r \varepsilon_\tau.$$

Donc, le second membre de la formule (15) est égal à

$$(-1)^r \sum_\tau \varepsilon_\tau \theta(X_{\tau(1)}, \ldots, X_{\tau(r)})\, \omega(X, X_{\tau(r+1)}, \ldots, X_{\tau(r+s-1)}) =$$
$$= (-1)^r (\theta \wedge (X \wedge \omega))(X_1, \ldots, X_{r+s-1}),$$

ce qui prouve la formule (14). □

On appelle *antidérivation* un opérateur linéaire D associant forme à forme, tel que pour toutes formes θ et ω

$$D(\theta \wedge \omega) = D\theta \wedge \omega + (-1)^r \theta \wedge D\omega,$$

où r est le degré de la forme θ. Dans cette terminologie, la proposition prouvée exprime que *pour tout champ de vecteurs X l'opérateur $i_X = X \lrcorner$ de multiplication intérieure par X est une antidérivation.*

* * *

Soit $f: \mathscr{X} \to \mathscr{Y}$ une application différentiable quelconque et soit ω une forme différentielle de degré $r \geqslant 0$ sur $\mathscr{Y}$. A tout point $p \in \mathscr{X}$ associons un tenseur antisymétrique $(f^*\omega)_p$ sur $\mathrm{T}_p\mathscr{X}$ prenant sur des vecteurs $A_1, \ldots, A_r \in \mathrm{T}_p\mathscr{X}$ la valeur

$$(f^*\omega)_p (A_1, \ldots, A_r) = \omega_q ((\mathrm{d}f)_p A_1, \ldots, (\mathrm{d}f)_p A_r),$$

où comme toujours $q = f(p)$ et

$$(\mathrm{d}f)_p : \mathrm{T}_p \mathscr{X} \to \mathrm{T}_p\mathscr{Y}$$

est la différentielle de l'application f en p. Si

$$(U, h) = (U, x^1, \ldots, x^n) \quad \text{et} \quad (V, k) = (V, y^1, \ldots, y^m)$$

sont des cartes des variétés $\mathscr{X}$ et $\mathscr{Y}$ telles que $fU \subset V$ et si

$$y^j = f^j(x^1, \ldots, x^n), \quad j = 1, \ldots, m,$$

sont des onctions définissant l'application f dans les cartes (U, h) et (V, k), alors

$$(\mathrm{d}f)_p \left(\frac{\partial}{\partial x^i}\right)_p = \left(\frac{\partial f^j}{\partial x^i}\right)_p \left(\frac{\partial}{\partial y^j}\right)_q, \quad i = 1, \ldots, n,$$

donc

$$(f^*\omega)_p \left(\left(\frac{\partial}{\partial x^{i_1}}\right)_p, \ldots, \left(\frac{\partial}{\partial x^{i_r}}\right)_p\right) =$$
$$= \left(\frac{\partial f^{j_1}}{\partial x^{i_1}}\right)_p \cdots \left(\frac{\partial f^{j_r}}{\partial x^{i_r}}\right)_p \omega_q \left(\left(\frac{\partial}{\partial y^{j_1}}\right)_q, \ldots, \left(\frac{\partial}{\partial y^{j_r}}\right)_q\right)$$

pour tous indices $i_1, \ldots, i_r = 1, \ldots, n$. Mais par définition

$$\omega_q \left(\left(\frac{\partial}{\partial y^{j_1}}\right)_q, \ldots, \left(\frac{\partial}{\partial y^{j_r}}\right)_q\right) = \omega_{j_1 \ldots j_r}(q) = (\omega_{j_1 \ldots j_r} \circ f)(p),$$

où $\omega_{j_1 \ldots j_r}$ sont les coefficients de la forme ω dans la carte (V, k). En introduisant de façon analogue les fonctions

$$(f^*\omega)_{i_1 \ldots i_r} : p \mapsto (f^*\omega)_p \left(\left(\frac{\partial}{\partial x^{i_1}}\right)_p, \ldots, \left(\frac{\partial}{\partial x^{i_r}}\right)_p \right),$$

on obtient donc pour tous indices $i_1, \ldots, i_r = 1, \ldots, n$ l'égalité

$$(16) \qquad (f^*\omega)_{i_1 \ldots i_r} = \frac{\partial f^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial f^{j_r}}{\partial x^{i_r}} (\omega_{j_1 \ldots j_r} \circ f),$$

sur U qui montre en particulier que *les fonctions* $(f^*\omega)_{i_1 \ldots i_r}$ *sont différentiables sur* U.

Donc, la formule

$$(17) \qquad f^*\omega = \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \ldots \sum (f^*\omega)_{i_1 \ldots i_r} \, dx^{i_1} \wedge \ldots \wedge dx^{i_r}$$

définit une forme différentielle $f^*\omega$ sur U.

La comparaison des définitions montre maintenant qu'en chaque point $p \in U$ cette forme prend la valeur $(f^*\omega)_p$. Ceci prouve que *la correspondance*

$$p \mapsto (f^*\omega)_p$$

définit une forme différentielle $f^*\omega$ *sur* $\mathcal{X}$. La forme $f^*\omega$ s'exprime par la formule (17) dans chaque carte $(U, x^1, \ldots, x^n)$ (i.e. les fonctions (16) sont les coefficients de cette forme dans la carte $(U, x^1, \ldots, x^n)$).

Définition 1. On dit de la forme $f^*\omega$ qu'elle se déduit de la forme ω moyennant un *transfert* par l'application différentiable f.

Il est clair que l'application

$$f^* : \Omega^r \mathcal{Y} \to \Omega^r \mathcal{X}, \qquad \omega \mapsto f^*\omega,$$

est linéaire et commute à la multiplication extérieure :

$$f^* (\theta \wedge \omega) = f^*\theta \wedge f^*\omega$$

pour toutes formes θ et ω sur $\mathcal{Y}$.

De plus, si $f : \mathcal{X} \to \mathcal{Y}$ et $g : \mathcal{Y} \to \mathcal{Z}$, alors

$$(g \circ f)^* = f^* \circ g^*,$$

et si $f = \mathrm{id} : \mathcal{X} \to \mathcal{X}$ est l'application identique, alors $f^* : \Omega^r \mathcal{X} \to \Omega^r \mathcal{X}$ est aussi une application identique.

Pour $r = 0$, dans le cas où la forme ω est une fonction différentiable $g : \mathcal{Y} \to \mathbb{R}$, on a

$$f^*g = g \circ f.$$

Lorsque f est un difféomorphisme φ, la construction de la forme $f^*\omega$ est un cas particulier de la construction générale de φ^*S de la leçon 17. (Comparer avec la remarque 1 de la leçon 17.)

Pour toute sous-variété $\mathscr{Y}$ de la variété $\mathscr{X}$ et toute injection $\iota : \mathscr{Y} \to \mathscr{X}$, l'application

$$\iota^* : \Omega^r\mathscr{X} \to \Omega^r\mathscr{Y}$$

est une *restriction* associant à la forme ω sur $\mathscr{X}$ la forme $\omega\,|_{\mathscr{Y}}$ sur $\mathscr{Y}$, telle que

$$(\omega\,|_{\mathscr{Y}})_p\,(A_1, \ldots, A_r) = \omega_p\,(A_1, \ldots, A_r)$$

en tout point $p \in \mathscr{Y}$ et pour tous vecteurs $A_1, \ldots, A_r \in \mathrm{T}_p\mathscr{Y}$ (où l'espace $\mathrm{T}_p\mathscr{Y}$ est de toute évidence traité comme un sous-espace de l'espace $\mathrm{T}_p\mathscr{X}$).

Remarque 3. La construction de la forme $f^*\omega$ se généralise immédiatement aux champs de tenseurs de type $(r, 0)$, $r \geqslant 0$. Pour un tel champ S sur $\mathscr{Y}$, le champ f^*S sur $\mathscr{X}$ est défini par la formule

$$(f^*S)_p\,(A_1, \ldots, A_r) = S_q\,((\mathrm{d}f)_p\,A_1, \ldots, (\mathrm{d}f)_p\,A_r),$$

où $p \in \mathscr{X}$, $q = f(p)$ et $A_1, \ldots, A_r \in \mathrm{T}_p\mathscr{X}$, et les composantes $(f^*S)_{i_1\ldots i_r}$ du champ f^*S s'expriment en fonction de celles $S_{i_1\ldots i_r}$ du champ S par la formule

$$(f^*S)_{i_1\ldots i_r} = \frac{\partial f^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial f^{j_r}}{\partial x^{i_r}}\,(S_{i_1\ldots i_r} \circ f).$$

Signalons que cette construction ne se généralise pas aux champs de type (r, s), avec $s > 0$ (si, bien entendu, f n'est pas un difféomorphisme).

LEÇON 19

Différentielle extérieure d'une forme différentielle. — Dérivée de Lie d'une forme différentielle.

On sait de la leçon 12 que toute fonction différentiable f sur une variété $\mathscr{X}$ définit en tout point $p \in \mathscr{X}$ un covecteur $(\mathrm{d}f)_p$ agissant à l'aide de la formule

$$(\mathrm{d}f)_p\, A = Af, \quad A \in \mathrm{T}_p\mathscr{X},$$

et, par suite, la forme différentielle linéaire

$$\mathrm{d}f\colon p \mapsto (\mathrm{d}f)_p,$$

appelée *différentielle* de la fonction f. Dans chaque carte $(U, x^1, \ldots, x^n)$ cette forme s'exprime par la formule

$$\mathrm{d}f = \frac{\partial f}{\partial x^i}\,\mathrm{d}x^i$$

et donc est une forme différentiable. En tant que morphisme de $\mathsf{F}\mathscr{X}$-modules $\mathfrak{a}\mathscr{X} \to \mathsf{F}\mathscr{X}$, la forme $\mathrm{d}f$ agit à l'aide de la formule

$$\mathrm{d}f\,(X) = Xf, \quad X \in \mathfrak{a}\mathscr{X}. \tag{1}$$

Il est clair que l'application

$$\mathrm{d}\colon \Omega^0\mathscr{X} \to \Omega^1\mathscr{X}, \quad f \mapsto \mathrm{d}f,$$

est linéaire et telle que

$$\mathrm{d}\,(fg) = \mathrm{d}f\cdot g + f\cdot \mathrm{d}g$$

pour toutes fonctions f et g.

L'application d se généralise de façon naturelle aux formes différentielles de tout degré.

Proposition 1. *Pour toute variété différentiable $\mathscr{X}$ et tout $r \geqslant 0$, il existe une seule application*

$$\mathrm{d}\colon \Omega^r\mathscr{X} \to \Omega^{r+1}\mathscr{X},$$

possédant les propriétés suivantes :

1° *L'application* d *est linéaire.*

2° *L'application* d *est une antidérivation, i.e. pour deux formes différentielles quelconques* θ *et* ω *on a*

$$\mathrm{d}(\theta \wedge \omega) = \mathrm{d}\theta \wedge \omega + (-1)^r \theta \wedge \mathrm{d}\omega,$$

où r *est le degré de la forme* θ.

3° *Pour toute application différentiable* $f : \mathscr{X} \to \mathscr{Y}$ *et toute forme* ω *sur* $\mathscr{Y}$, *on a l'égalité*

$$\mathrm{d}f^* \omega = f^* \mathrm{d}\omega.$$

4° *Pour toute fonction* $f \in \Omega^0 \mathscr{X}$, *la forme* $\mathrm{d}f$ *est sa différentielle* (1).

5° *Si* $\omega = \mathrm{d}f$, *où* $f \in \Omega^0 \mathscr{X}$, *alors*

$$\mathrm{d}\omega = 0.$$

D é m o n s t r a t i o n. Comme toujours en pareille situation, on commencera par prouver l'unicité.

Soit $(U, h) = (U, x^1, \ldots, x^n)$ une carte quelconque de $\mathscr{X}$, et soit

$$\omega|_U = \sum_{1 \leqslant i_1 \ldots < i_r \leqslant n} \ldots \sum \omega_{i_1 \ldots i_r} \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_r}.$$

En désignant la forme

$$\mathrm{d}(\omega|_U) = \mathrm{d}\omega|_U$$

par $\mathrm{d}\omega_U$, on trouve immédiatement en vertu des propriétés 1° à 5° que

$$(2) \qquad \mathrm{d}\omega_U = \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \ldots \sum \mathrm{d}\omega_{i_1 \ldots i_r} \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_r} =$$

$$= \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \ldots \sum \frac{\partial \omega_{i_1 \ldots i_r}}{\partial x^i} \mathrm{d}x^i \wedge \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_r}.$$

Donc, la forme $\mathrm{d}\omega_U = \mathrm{d}\omega|_U$ et par suite — puisque le voisinage de coordonnées U est arbitraire — la forme $\mathrm{d}\omega$ sont définies de façon unique par la forme ω. Ceci exprime que l'application d est unique.

Pour prouver son existence, définissons sur chaque voisinage de coordonnées U une forme $\mathrm{d}\omega_U$ à l'aide de la formule (2). Si $(U, x^1, \ldots, x^n)$ et $(U', x^{1'}, \ldots, x^{n'})$ sont deux cartes de la variété $\mathscr{X}$ et si

$$\omega = \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \ldots \sum \omega_{i_1 \ldots i_r} \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_r} \quad \text{sur } U$$

et

$$\omega = \sum_{1 \leqslant i'_1 < \ldots < i'_r \leqslant n} \ldots \sum \omega_{i'_1 \ldots i'_r} \mathrm{d}x^{i'_1} \wedge \ldots \wedge \mathrm{d}x^{i'_r} \quad \text{sur } U,$$

alors

$$\omega_{i_1\ldots i_r}=\frac{\partial x^{i'_1}}{\partial x^{i_1}}\cdots\frac{\partial x^{i'_r}}{\partial x^{i_r}}\omega_{i'_1\ldots i'_r}\quad\text{sur } U\cap U',$$

donc,

$$\frac{\partial\omega_{i_1\ldots i_r}}{\partial x^i}\sum_{k=1}^{r}\frac{\partial x^{i'_1}}{\partial x^{i_1}}\cdots\frac{\partial^2 x^{i'_k}}{\partial x^i\,\partial x^{i_k}}\cdots\frac{\partial x^{i'_r}}{\partial x^{i_r}}\omega_{i'_1\ldots i'_r}+\frac{\partial x^{i'_1}}{\partial x^{i_1}}\cdots$$

$$\cdots\frac{\partial x^{i'_r}}{\partial x^{i_r}}\frac{\partial\omega_{i'_1\ldots i'_r}}{\partial x^i}.$$

D'autre part, pour toute fonction f,

$$(3)\quad \frac{\partial^2 f}{\partial x^i\,\partial x^j}\,\mathrm{d}x^i\wedge\mathrm{d}x^j=\sum_{i<j}\frac{\partial^2 f}{\partial x^i\,\partial x^j}\,\mathrm{d}x^i\wedge\mathrm{d}x^j+$$

$$+\sum_{j<i}\frac{\partial^2 f}{\partial x^i\,\partial x^j}\,\mathrm{d}x^i\wedge\mathrm{d}x^j=\sum_{i<j}\frac{\partial^2 f}{\partial x^i\,\partial x^j}\,\mathrm{d}x^i\wedge\mathrm{d}x^j+$$

$$+\sum_{i<j}\frac{\partial^2 f}{\partial x^j\,\partial x^i}\,\mathrm{d}x^j\wedge\mathrm{d}x^i=\sum_{i<j}\left(\frac{\partial^2 f}{\partial x^i\,\partial x^j}-\frac{\partial^2 f}{\partial x^j\,\partial x^i}\right)\mathrm{d}x^i\wedge\mathrm{d}x^j=0.$$

car $\mathrm{d}x^i\wedge\mathrm{d}x^j=-\mathrm{d}x^j\wedge\mathrm{d}x^i$ et les dérivées partielles secondes ne dépendent pas de l'ordre de dérivation. Donc, en particulier pour tout $k=1,\ldots,r$,

$$\frac{\partial^2 x^{i'_k}}{\partial x^i\,\partial x^{i_k}}\,\mathrm{d}x^i\wedge\mathrm{d}x^{i_k}=0,$$

et par suite

$$\sum_{k=1}^{r}\frac{\partial x^{i'_1}}{\partial x^{i_1}}\cdots\frac{\partial^2 x^{i'_k}}{\partial x^i\,\partial x^{i_k}}\cdots\frac{\partial x^{i'_r}}{\partial x^{i_r}}\omega_{i'_1\ldots i'_r}\,\mathrm{d}x^i\wedge\mathrm{d}x^{i_1}\wedge\cdots$$

$$\cdots\wedge\mathrm{d}x^{i_k}\wedge\cdots\wedge\mathrm{d}x^{i_r}=0.$$

Donc, sur $U\cap U'$,

$$\frac{\partial\omega_{i_1\ldots i_r}}{\partial x^i}\,\mathrm{d}x^i\wedge\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_r}=$$

$$=\frac{\partial x^{i'_1}}{\partial x^{i_1}}\cdots\frac{\partial x^{i'_r}}{\partial x^{i_1}}\frac{\partial\omega_{i'_1\ldots i'_r}}{\partial x^r}\,\mathrm{d}x^i\wedge\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_r}=$$

$$=\left(\frac{\partial\omega_{i'_1\ldots i'_r}}{\partial x^l}\,\mathrm{d}x^i\right)\wedge\left(\frac{\partial x^{i'_1}}{\partial x^{i_1}}\,\mathrm{d}x^{i_1}\right)\wedge\cdots\wedge\left(\frac{\partial x^{i'_r}}{\partial x^{i_r}}\,\mathrm{d}x^{i_r}\right)=$$

$$=\mathrm{d}\omega_{i'_1\ldots i'_r}\wedge\mathrm{d}x^{i'_1}\wedge\cdots\wedge\mathrm{d}x^{i'_r}=\frac{\partial\omega_{i'_1\ldots i'_r}}{\partial x^{i'}}\,\mathrm{d}x^{i'}\wedge\mathrm{d}x^{i'_1}\wedge\cdots\wedge\mathrm{d}x^{i'_r},$$

et, par suite,

$$\sum_{1\leqslant i_1<\ldots<i_r\leqslant n} \cdots \sum \frac{\partial\omega_{i_1\ldots i_r}}{\partial x^i} dx^i \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_r} =$$

$$= \sum_{1\leqslant i_1<\ldots<i_r\leqslant n} \cdots \sum \frac{\partial\omega_{i'_1\ldots i'_r}}{\partial x^{i'}} dx^{i'} \wedge dx^{i'_1} \wedge \ldots \wedge dx^{i'_r}$$

sur $U \cap U'$, i.e.

$$d\omega_U = d\omega_{U'} \quad \text{sur } U \cap U'.$$

Les formes $d\omega_U$ sont donc compatibles sur les intersections et, par suite, engendrent une forme différentielle $d\omega$ de degré $r+1$ sur $\mathscr{X}$, telle que

$$d\omega\,|_U = d\omega_U$$

pour tout voisinage de coordonnées U.

Nous avons ainsi construit l'application

$$d : \Omega^r\mathscr{X} \to \Omega^{r+1}\mathscr{X}.$$

Il est clair que cette application possède les propriétés 1° et 4°. De plus, en vertu de la formule (3), pour toute fonction $f \in \mathsf{F}\mathscr{X}$ on a dans chaque carte $(U, x^1, \ldots, x^n)$

$$d(df) = d\left(\frac{\partial f}{\partial x^j} dx^j\right) = \frac{\partial^2 f}{\partial x^i\,\partial x^j} dx^i \wedge dx^j = 0,$$

ce qui prouve la propriété 5°.

Il suffit de toute évidence de vérifier la propriété 2° pour une carte arbitraire $(U, x^1, \ldots, x^n)$. En outre, en raison de la linéarité de l'opérateur d, il suffit de vérifier cette propriété pour des formes « monômes »

$$\theta = f\,dx^\alpha \quad \text{et} \quad \omega = g\,dx^\beta,$$

où

$$dx^\alpha = dx^{i_1} \wedge \ldots \wedge dx^{i_r} \quad \text{et} \quad dx^\beta = dx^{j_1} \wedge \ldots \wedge dx^{j_s}.$$

Mais pour de telles formes

$$\begin{aligned} d(\theta \wedge \omega) &= d(fg \wedge dx^\alpha \wedge dx^\beta) = d(fg) \wedge dx^\alpha \wedge dx^\beta = \\ &= (df \cdot g + f \cdot dg) \wedge dx^\alpha \wedge dx^\beta = \\ &= df \wedge dx^\alpha \wedge g\,dx^\beta + (-1)^r (f \wedge dx^\alpha) \wedge (dg \wedge dx^\beta) = \\ &= d\theta \wedge \omega + (-1)^r \theta \wedge d\omega, \end{aligned}$$

ce qui prouve la propriété 2°.

De façon analogue, il suffit de vérifier la propriété 3° seulement pour des voisinages de coordonnées (U, h) et (V, k) tels que $fU \subset V$.

Mais dans ce cas, si

$$\omega = \sum_{1 \leqslant j_1 < \ldots < j_r \leqslant n} \ldots \sum \omega_{j_1 \ldots j_r} \, dy^{j_1} \wedge \ldots \wedge dy^{j_r} \quad \text{sur} \quad V,$$

alors (cf. formules (16) et (17) de la leçon 18)

$$d(f^*\omega) = \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \ldots \sum d(f^*\omega)_{i_1 \ldots i_r} \, dx^{i_1} \wedge \ldots \wedge dx^{i_r},$$

où

$$d(f^*\omega)_{i_1 \ldots i_r} = d\left(\frac{\partial f^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial f^{j_r}}{\partial x^{i_r}} (\omega_{j_1 \ldots j_r} \circ f)\right) =$$

$$= \sum_{s=1}^{r} \frac{\partial f^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial^2 f^{j_s}}{\partial x^i \, \partial x^{i_s}} \cdots \frac{\partial f^{j_r}}{\partial x^{i_r}} (\omega_{j_1 \ldots j_r} \circ f) \, dx^i +$$

$$+ \frac{\partial f^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial f^{j_r}}{\partial x^{i_r}} \left(\frac{\partial \omega_{j_1 \ldots j_r}}{\partial y^j} \circ f\right) \frac{\partial f^j}{\partial x^i} \, dx^i,$$

et donc (cf. ci-dessus les calculs analogues impliqués par la construction de l'opérateur d),

$$d(f^*\omega) = \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \ldots \sum \frac{\partial f^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial f^{j_r}}{\partial x^{i_r}} \left(\frac{\partial \omega_{j_1 \ldots j_r}}{\partial y^j} \circ f\right) \frac{\partial f^j}{\partial x^i} \, dx^i \wedge$$

$$\wedge \, dx^{i_1} \wedge \ldots \wedge dx^{i_r}.$$

D'autre part,

$$d\omega = \sum_{1 \leqslant j_1 < \ldots < j_r \leqslant n} \ldots \sum \frac{\partial \omega_{j_1 \ldots j_r}}{\partial y^j} \, dy^j \wedge dy^{j_1} \wedge \ldots \wedge dy^{j_r}$$

et

$$f^*(d\omega) =$$

$$= \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \ldots \sum \left(\frac{\partial \omega_{j_1 \ldots j_r}}{\partial y^j} \circ f\right) f^* \, dy^j \wedge f^* \, dy^{j_1} \wedge \ldots \wedge f^* \, dy^{j_r} =$$

$$= \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \ldots \sum \frac{\partial f^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial f^{j_r}}{\partial x^{i_r}} \left(\frac{\partial \omega_{j_1 \ldots j_r}}{\partial y^j} \circ f\right) \frac{\partial f^j}{\partial x^i} \, dx^i \wedge$$

$$\wedge \, dx^{i_1} \wedge \ldots \wedge dx^{i_r}.$$

Donc, $d(f^*\omega) = f^*(d\omega)$. □

Définition 1. La forme $d\omega$ s'appelle *différentielle extérieure* de la forme ω.

Pour exprimer les coefficients $(d\omega)_{j_1 \ldots j_{r+1}}$ de la forme $d\omega$ dans la carte $(U, h) = (U, x^1, \ldots, x^n)$ en fonction de ceux $\omega_{i_1 \ldots i_r}$

de la forme ω, on remarquera tout d'abord que dans la somme (2) on peut se limiter à la sommation sur les indices distincts $i, i_1, \ldots, i_r$, puisque, si l'indice i est confondu avec l'un des indices $i_1, \ldots, i_r$ (qui par hypothèse sont tous distincts), le terme correspondant de la somme (2) sera nul. Si les indices $i, i_1, \ldots, i_r$ sont tous distincts, rangeons-les par ordre de grandeur croissante et désignons-les par $j_1, \ldots, j_{r+1}$. Donc, si $i = j_a$, où $a = 1, \ldots, r+1$, alors

$$i_1 = j_1, \ldots, \quad i_{a-1} = j_{a-1}, \quad i_a = j_{a+1}, \ldots, \quad i_r = j_{r+1}$$

(on rappelle que par hypothèse $i_1 < \ldots < i_r$). Ceci étant,

$$dx^i \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_r} = (-1)^{a-1} dx^{j_1} \wedge \ldots \wedge dx^{j_a} \wedge \ldots \wedge dx^{j_{r+1}}$$

et

$$\frac{\partial \omega_{i_1 \ldots i_r}}{\partial x^i} = \frac{\partial \omega_{j_1 \ldots \hat{j}_a \ldots j_{r+1}}}{\partial x^{j_a}},$$

où le signe ^ indique que l'indice correspondant doit être omis. Puisque l'indice i peut être l'un quelconque des indices croissants $j_1, \ldots, j_{r+1}$, on en déduit que

$$\sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \cdots \sum \frac{\partial \omega_{i_1 \ldots i_r}}{\partial x^i} dx^i \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_r} =$$

$$= \sum_{1 \leqslant j_1 < \ldots < j_{r+1} \leqslant n} \cdots \sum \left(\sum_{a=1}^{r+1} (-1)^{a+1} \frac{\partial \omega_{j_1 \ldots \hat{j}_a \ldots j_{r+1}}}{\partial x^{j_a}} \right) dx^{j_1} \wedge \ldots \wedge dx^{j_{r+1}},$$

i.e. que

$$(4) \qquad (d\omega)_{j_1 \ldots j_{r+1}} = \sum_{a=1}^{r+1} (-1)^{a+1} \frac{\partial \omega_{j_1 \ldots \hat{j}_a \ldots j_{r+1}}}{\partial x^{j_a}}.$$

Par exemple, pour la forme linéaire $\alpha = \alpha_i dx^i$,

$$(d\alpha)_{ij} = \frac{\partial \alpha_j}{\partial x^i} - \frac{\partial \alpha_i}{\partial x^j}, \quad i < j.$$

Proposition 2. *Traitée comme une fonctionnelle* $\mathsf{F}\mathscr{X}$*-multilinéaire de champs de vecteurs, la forme* $d\omega$ *est définie par la formule*

$$(5) \quad (d\omega)(X_1, \ldots, X_{r+1}) = \sum_{a=1}^{r+1} (-1)^{a+1} X_a \omega(X_1, \ldots, \hat{X}_a, \ldots, X_{r+1}) +$$

$$+ \sum_{a=1}^{r} \sum_{b=a+1}^{r+1} (-1)^{a+b} \omega([X_a, X_b], X_1, \ldots, \hat{X}_a, \ldots, \hat{X}_b, \ldots, X_{r+1}),$$

où le signe ^ indique que le champ correspondant doit être omis.

D é m o n s t r a t i o n. Soit $\theta(X_1, \ldots, X_{r+1})$ le second membre de la formule (5). Il est clair que

$$\theta(X_1 + X_1', X_2, \ldots, X_{r+1}) = \\ = \theta(X_1, X_2, \ldots, X_{r+1}) + \theta(X_1', X_2, \ldots, X_{r+1})$$

pour tous champs $X_1, X_1', X_2, \ldots, X_{r+1}$. De plus, pour toute fonction $f \in \mathsf{F}\mathscr{X}$,

$$\theta(fX_1, X_2, \ldots, X_{r+1}) = fX_1\omega(X_2, \ldots, X_{r+1}) + \\ + \sum_{a=2}^{r+1}(-1)^{a+1}X_a\omega(fX_1, \ldots, \hat{X}_a, \ldots, X_{r+1}) + \\ + \sum_{b=2}^{r+1}(-1)^{b+1}\omega([fX_1, X_b], X_2, \ldots, \hat{X}_b, \ldots, X_{r+1}) + \\ + \sum_{a=2}^{r}\sum_{b=a+1}^{r+1}(-1)^{a+b}\omega([X_a, X_b], fX_1, \ldots \\ \ldots, \hat{X}_a, \ldots, \hat{X}_b, \ldots, X_{r+1}) = fX_1\omega(X_2, \ldots, X_{r+1}) + \\ + \sum_{a=2}^{r+1}(-1)^{a+1}X_a(f\omega(X_1, \ldots, \hat{X}_a, \ldots, X_{r+1})) + \\ + \sum_{b=2}^{r+1}(-1)^{b+1}[f\omega([X_1, X_b], X_2, \ldots, \hat{X}_b, \ldots, X_{r+1}) - \\ - X_bf\cdot\omega(X_1, \ldots, \hat{X}_b, \ldots, X_{r+1})] + \\ + \sum_{a=2}^{r}\sum_{b=a+1}^{r+1}(-1)^{a+b}f\omega([X_a, X_b], X_1, \ldots \\ \ldots, \hat{X}_a, \ldots, \hat{X}_b, \ldots, X_{r+1}) = fX_1\omega(X_2, \ldots, X_{r+1}) + \\ + \sum_{a=2}^{r+1}(-1)^{a+1}X_af\cdot\omega(X_1, \ldots, \hat{X}_a, \ldots, X_{r+1}) + \\ + \sum_{a=2}^{r+1}(-1)^{a+1}fX_a\omega(X_1, \ldots, \hat{X}_a, \ldots, X_{r+1}) + \\ + \sum_{b=2}^{r+1}(-1)^{b+1}f\omega([X_1, X_b], X_2, \ldots, \hat{X}_b, \ldots, X_{r+1}) - \\ - \sum_{b=2}^{r+1}(-1)^{b+1}X_bf\cdot\omega(X_1, \ldots, \hat{X}_b, \ldots, X_{r+1}) + \\ + \sum_{a=2}^{r}\sum_{b=a+1}^{r+1}(-1)^{a+b}f\omega([X_a, X_b], X_1, \ldots, \hat{X}_a, \ldots, \hat{X}_b, \ldots, X_{r+1})$$

(on s'est servi de la relation $[fX_1, X_b] = f[X_1, X_b] - X_b f \cdot X_1$; cf. formule (23) de la leçon 16) et, par suite,

$$\theta(fX_1, X_2, \ldots, X_{r+1}) - f\theta(X_1, X_2, \ldots, X_{r+1}) =$$
$$= \sum_{a=2}^{r+1} (-1)^{a+1} X_a f \cdot \omega(X_1, \ldots, \hat{X}_a, \ldots, X_{r+1}) -$$
$$- \sum_{b=2}^{r+1} (-1)^{b+1} X_b f \cdot \omega(X_1, \ldots, \hat{X}_b, \ldots, X_{r+1}) = 0.$$

Ceci prouve que *le second membre de la formule* (5) *est* $\mathsf{F}\mathscr{X}$*-linéaire par rapport à* X_1.

On prouve de façon analogue (le lecteur est invité à effectuer les calculs nécessaires) qu'elle est $\mathsf{F}\mathscr{X}$-linéaire par rapport aux autres champs $X_2, \ldots, X_{r+1}$. (Une autre démonstration un peu plus facile consiste à prouver que $\theta(X_1, \ldots, X_{r+1})$ est antisymétrique par rapport à $X_1, \ldots, X_{r+1}$.)

Donc, il suffit — dans chaque carte $(U, x^1, \ldots, x^n)$ — de vérifier la formule (5) pour les seuls champs de vecteurs de base

$$X_1 = \frac{\partial}{\partial x^{j_1}}, \ldots, X_{r+1} = \frac{\partial}{\partial x^{j_{r+1}}}.$$

En calculant les deux membres de la formule (5) sur ces champs, on obtient au premier le coefficient $(d\omega)_{j_1 \ldots j_{r+1}}$ de la forme $d\omega$ et, au second, la somme

$$\sum_{a=1}^{r+1} (-1)^{a+1} \frac{\partial}{\partial x^{j_a}} \omega_{j_1 \ldots \hat{j} \ldots j_{r+1}}$$

(la deuxième somme disparaît, puisque $\left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right] = 0$ pour tous i et j). En vertu de la formule (4), ceci prouve la proposition 2.□

Remarque 1. On pourrait tenter de prouver la proposition en vérifiant que la forme $d\omega$ définie par la formule (5) possède les propriétés 1° à 5° de la proposition 1. La propriété 1° (la linéarité de l'application d) et la propriété 3° (la commutativité à l'application de transfert f^*) sont évidentes. Pour $r = 0$ (et $\omega = f$) la formule (5) se transforme en la formule (1), ce qui prouve la propriété 4°. Pour $r = 1$ la formule (5) devient

$$d\omega(X, Y) = X\omega(Y) - Y\omega(X) - \omega([X, Y]). \tag{6}$$

Donc, si $\omega = df$ (et, par suite, $\omega(Z) = Zf$ pour tout champ Z), alors

$$d\omega(X, Y) = XYf - YXf - [X, Y]f = 0,$$

ce qui prouve la propriété 5°. La vérification de la propriété 2° (qui s'effectue à l'aide de la formule (13) de la leçon 18) implique des calculs peu compliqués mais assez laborieux sur des sommes doubles.

Exercice 1. Effectuer ces calculs (et par la même occasion prouver de nouveau la proposition 2).

* * *

Il s'avère que pour les formes différentielles la dérivation de Lie $\pounds_X$ s'exprime en fonction du produit intérieur par le champ X et de l'opérateur d.

Proposition 3. *Pour tout champ de vecteurs X et toute forme différentielle ω on a l'égalité*

$$\pounds_X\omega = X \lrcorner\, d\omega + d\,(X \lrcorner\, \omega). \tag{7}$$

D é m o n s t r a t i o n. Si une forme ω est de degré $r = 0$, i.e. est une fonction f, on sait (cf. formule (14) de la leçon 17) que le premier membre de la formule (7) est égal à Xf. Au second membre il ne reste alors que le premier terme $X \lrcorner\, df = df\,(X)$. Ceci prouve la formule (7) pour $r = 0$ puisque $df\,(X) = Xf$.

De la propriété 3° de l'opérateur d il s'ensuit immédiatement que *cet opérateur commute à l'opérateur* $\pounds_X$, i.e.

$$d\pounds_X\omega = \pounds_X\, d\omega \tag{8}$$

pour toute forme ω. En particulier,

$$\pounds_X\, df = d\pounds_X f = d\,(Xf)$$

pour toute fonction f.

D'autre part, si $\omega = df$, alors

$$X \lrcorner\, d\omega + d\,(X \lrcorner\, \omega) = d\,(X \lrcorner\, df) = d\,(Xf).$$

Donc, la formule (7) est valable pour $\omega = df$.

Le produit extérieur des formes s'obtenant par alternation de leur produit tensoriel, la multiplication extérieure est justiciable d'une formule analogue à la formule (13) de la leçon (17), i.e.

$$\pounds_X\,(\theta \wedge \omega) = \pounds_X\theta \wedge \omega + \theta \wedge \pounds_X\omega \tag{9}$$

pour toutes formes θ et ω. En vertu de la définition 1 de la leçon 16, ceci exprime que *l'opérateur $\pounds_X$ est une dérivation de l'algèbre des formes sur $\mathcal{X}$.*

On sait au contraire que les opérateurs d et $X \lrcorner$ sont des antidérivations. Il est aisé de voir que pour deux antidérivations quelconques D_1 et D_2 dont le degré varie de ± 1 (ou de façon plus générale d'un nombre impair quelconque) *l'opérateur $D_1D_2 + D_2D_1$ est*

une dérivation. (En effet,

$$\begin{aligned}(D_1D_2 + D_2D_1)(\theta\wedge\omega) &= D_1(D_2\theta\wedge\omega + (-1)^r\theta\wedge D_2\omega) + \\ &+ D_2(D_1\theta\wedge\omega + (-1)^r\theta\wedge D_1\omega) = D_1D_2\theta\wedge\omega + \\ &+ (-1)^{r+1}D_2\theta\wedge D_1\omega + (-1)^r D_1\theta\wedge D_2\omega + (-1)^{r+r}\theta\wedge \\ &\wedge D_1D_2\omega + D_2D_1\theta\wedge\omega + (-1)^{r\pm1}D_1\theta\wedge D_2\omega + \\ &+ (-1)^r D_2\theta\wedge D_1\omega + (-1)^{r+r}\theta\wedge D_2D_1\omega = \\ &= (D_1D_2 + D_2D_1)\theta\wedge\omega + \theta\wedge(D_1D_2 + D_2D_1)\omega\end{aligned}$$

pour toutes formes θ et ω.) En particulier, l'*opérateur*

$$\omega \longmapsto X \lrcorner\, d\omega + d(X \lrcorner\, \omega) \tag{10}$$

du second membre de la relation (7) *est une dérivation.*

Puisque sur tout voisinage de coordonnées U l'algèbre des formes est engendrée par les fonctions et les formes dx^i sur lesquelles, comme nous l'avons déjà vu, les opérateurs $\pounds_X$ et (10) sont confondus et puisque le fait que deux dérivations d'une algèbre agissent de la même manière sur les génératrices entraîne de toute évidence que ces dérivations sont partout confondues, la formule (7) a lieu sur chaque voisinage de coordonnées U. Elle est donc valable sur la variété $\mathscr{X}$ tout entière. □

La relation (7) peut se mettre sous la forme opératorielle

$$\pounds_X = i_X \circ d + d \circ i_X.$$

Exercice 2. Montrer que

$$[\pounds_X, i_Y] = i_{[X,Y]}$$

pour tous champs de vecteurs X, Y. [N o t a. Les deux membres de cette formule sont des antidérivations abaissant le degré de —1. De plus, ils sont nuls sur les fonctions et agissent de la même façon sur les formes df.]

Exercice 3. Montrer que

$$[\pounds_X, \pounds_Y] = \pounds_{[X,Y]}$$

pour tous champs de tenseurs. [N o t a. Les deux membres de cette formule sont des dérivations de l'algèbre des champs de tenseurs agissant de la même façon sur les fonctions, les champs de vecteurs et les formes différentielles linéaires.]

LEÇON 20

Complexe de de Rham et groupes de cohomologie d'une variété différentiable. — Groupe $H^0\mathscr{X}$. — Lemme de Poincaré. — Groupe H^1S^2. — Groupe H^1S^1. — Calcul du groupe H^1S^1 à l'aide d'intégrales. — Groupe H^2S^2. — Groupes H^1S^n pour $n \geqslant 2$. — Groupes H^mS^n, $m < n$. — Groupes H^nS^n.

L'égalité

$$d\omega = \theta \tag{1}$$

où θ est une forme donnée et ω, une forme inconnue, se représente dans toute carte par un système d'équations différentielles sur les coefficients de la forme ω. Ces équations étant partout étudiées en mathématiques et en physique, on se penchera dans cette leçon et dans les suivantes sur les conditions d'existence de leurs solutions.

* * *

Proposition 1. *Pour toute forme différentielle ω, on a*

$$dd\omega = 0.$$

D é m o n s t r a t i o n. Il suffit de vérifier cette égalité dans une carte quelconque et seulement pour les formes $\omega = f\,dx^\alpha$ où $dx^\alpha = dx^{i_1} \wedge \ldots \wedge dx^{i_r}$. Mais par définition,

$$d(f\,dx^\alpha) = df \wedge dx^\alpha$$

et, en vertu des propriétés 2° et 5° de l'opérateur d (cf. proposition 1 de la leçon 19),

$$dd(f\,dx^\alpha) = d(df \wedge dx^\alpha) = ddf \wedge dx^\alpha - df \wedge ddx^\alpha = 0. \quad \square$$

La proposition 1 exprime que l'égalité $d\theta = 0$ est une c o n d i t i o n n é c e s s a i r e d'existence de la solution de l'équation (1). L'établissement de la suffisance de cette condition implique une étude générale des relations liant les formes $d\omega$ et les formes θ telles que $d\theta = 0$.

Définition 1. La suite

$$C^\bullet : C^0 \xrightarrow{d^0} C^1 \xrightarrow{d^1} \ldots \xrightarrow{d^{m-1}} C^m \xrightarrow{d^m} C^{m+1} \rightarrow \ldots$$

de groupes (ou en particulier d'espaces vectoriels) et d'homomorphismes s'appelle *complexe de cochaînes* (l'origine de ce terme apparaîtra plus bas à la leçon 29) si

$$\mathrm{d}^m \circ \mathrm{d}^{m-1} = 0 \quad \text{pour tout } m \geqslant 1. \tag{2}$$

Généralement au lieu de d^m on écrit simplement d.

Le noyau Ker d^m de l'homomorphisme d^m est désigné par le symbole $Z^m C^\cdot$, et ses éléments s'appellent *cocycles de degré m*. L'image Im d^{m-1} de l'homomorphisme d^{m-1} est désignée par le symbole $B^m C^\cdot$, et ses éléments s'appellent *cobords de degré m*. (Pour $m = 0$ on convient que $B^0 C^\cdot = 0$.) Puisque $B^m C^\cdot \subset Z^m C^\cdot$ en vertu de la condition (2), est défini le groupe quotient

$$H^m C^\cdot = Z^m C^\cdot / B^m C^\cdot. \tag{3}$$

Ce groupe quotient s'appelle *m-ième groupe de cohomologie* du complexe $C^\cdot$, et ses éléments, *classes de cohomologie*. Les cocycles appartenant à une même classe de cohomologie, i.e. dont la différence est un cobord, sont dits *cohomologues*.

En vertu de la proposition 1, les espaces vectoriels $\Omega^m \mathscr{X}$ et les différentielles extérieures $\mathrm{d}: \Omega^m \mathscr{X} \to \Omega^{m+1} \mathscr{X}$ forment un complexe de cochaînes

$$\Omega^\cdot \mathscr{X} : \Omega^0 \mathscr{X} \xrightarrow{\mathrm{d}} \Omega^1 \mathscr{X} \to \ldots \to \Omega^m \mathscr{X} \xrightarrow{\mathrm{d}} \Omega^{m+1} \mathscr{X} \to \ldots$$

Ce complexe s'appelle *complexe de de Rham* des formes différentielles de la variété différentiable $\mathscr{X}$. Ses groupes de cohomologie sont désignés par les symboles $H^m \mathscr{X}$ et appelés *groupes de cohomologie de de Rham* de la variété $\mathscr{X}$. (Remarquons qu'en fait ce sont des espaces vectoriels, mais le terme « espace vectoriel de cohomologie » n'est pas usité.)

Les cobords d'un complexe de de Rham, i.e. les formes $\mathrm{d}\omega$, s'appellent *différentielles exactes*, ou plus brièvement *formes exactes*.

Les cocycles d'un complexe de de Rham, i.e. les formes ω telles que $\mathrm{d}\omega = 0$, s'appellent *formes fermées*. La classe de cohomologie d'une forme fermée ω sera désignée par le symbole $[\omega]$.

Donc, une condition nécessaire d'existence de la solution de l'équation (1) est la fermeture de la forme θ, et une condition suffisante, son exactitude. En particulier, l'équation (1) admet une solution pour toute forme fermée θ de degré m si et seulement si $H^m \mathscr{X} = 0$. [Ces assertions sont de toute évidence des tautologies. Elles deviendront consistantes quand nous apprendrons géométriquement à calculer le groupe $H^m \mathscr{X}$.]

Soient $C^\cdot = \{C^m, \mathrm{d}^m\}$ et $D^\cdot = \{D^m, \delta^m\}$ deux complexes de cochaînes.

Définition 2. On appelle *application de cochaînes*

$$\varphi^\bullet : C^\cdot \to D^\bullet$$

d'un complexe $C^\cdot$ dans un complexe $D^\cdot$ une suite d'applications

$$\varphi^m : C^m \to D^m, \quad m \geqslant 0,$$

telle que

$$\varphi^{m+1} \circ d^m = \delta^m \circ \varphi^m$$

pour tout $m \geqslant 0$, i.e. telle que le diagramme

$$\begin{tikzcd}
C^0 \arrow[r, "d^0"] \arrow[d, "\varphi^0"'] & C^1 \arrow[r] \arrow[d, "\varphi^1"'] & \dots \arrow[r] & C^m \arrow[r, "d^m"] \arrow[d, "\varphi^m"'] & C^{m+1} \arrow[r] \arrow[d, "\varphi^{m+1}"] & \dots \\
D^0 \arrow[r, "\delta^0"] & D^1 \arrow[r] & \dots \arrow[r] & D^m \arrow[r, "\delta^m"] & D^{m+1} \arrow[r] & \dots
\end{tikzcd}$$

soit commutatif. Au lieu de $\varphi^\cdot$ et φ^m on écrit généralement tout simplement φ.

Il est clair que toute application de cochaînes $\varphi : C^\cdot \to D^\cdot$ envoie le groupe $Z^m C^\cdot$ dans le groupe $Z^m D^\cdot$, et chaque groupe $B^m C^\cdot$ dans le groupe $B^m D^\cdot$. Donc, pour tout $m \geqslant 0$, il induit une application

$$\varphi^* : H^m C^\cdot \to H^m D^\cdot.$$

Ceci étant, pour toutes applications de cochaînes $\varphi : C^\cdot \to D^\cdot$ et $\psi : D^\cdot \to E^\cdot$ on a

$$(\psi \circ \varphi)^* = \psi^* \circ \varphi^*.$$

De plus, si $\varphi = \mathrm{id}$, alors $\varphi^* = \mathrm{id}$. [Ces propriétés expriment la *fonctorialité* de la correspondance $\varphi \mapsto \varphi^*$. Nous les avons rencontrées à plusieurs reprises.]

La propriété 3° de l'opérateur d figurant dans la proposition 1 de la leçon 19 exprime que *pour toute application différentiable* $f : \mathscr{X} \to \mathscr{Y}$ *les homomorphismes* $f^* : \Omega^m \mathscr{Y} \to \Omega^m \mathscr{X}$ *forment une application de cochaînes*

$$f^\cdot : \Omega^\cdot \mathscr{Y} \to \Omega^\cdot \mathscr{X}$$

de complexes de de Rham. Les homomorphismes de groupes de cohomologie induits par cette application de cochaînes sont désignés par le même symbole f^* (aucune confusion n'est à craindre) et s'appellent *homomorphismes induits par l'application* f.

Donc, par définition,

$$f^* [\omega] = [f^* \omega]$$

pour toute forme fermée ω de degré m sur la variété $\mathscr{Y}$.

Il est clair que si $f = \mathrm{id} : \mathscr{X} \to \mathscr{X}$ est l'application identique, alors $f^* : H^m \mathscr{X} \to H^m \mathscr{X}$ est aussi une application identique, et pour toutes applications différentiables $f : \mathscr{X} \to \mathscr{Y}$ et $g : \mathscr{Y} \to \mathfrak{Z}$

on a l'égalité

$$(g \circ f)^* = f^* \circ g^* : H^m\mathscr{X} \to H^m\mathscr{Z}.$$

[La correspondance $f \mapsto f^*$ possède la propriété de fonctorialité.]

Dans le cas particulier où $\mathscr{X}$ est une sous-variété de la variété $\mathscr{Y}$, et f est une injection de $\mathscr{X}$ dans $\mathscr{Y}$, la classe de cohomologie $f^*[\omega]$ est désignée par le symbole $[\omega]|_{\mathscr{X}}$ et s'appelle *restriction* de la classe $[\omega]$ à $\mathscr{X}$. On a donc par définition

$$[\omega]|_{\mathscr{X}} = [\omega|_{\mathscr{X}}]$$

pour toute classe de cohomologie $[\omega] \in H^m\mathscr{Y}$.

La dimension

$$h^m\mathscr{X} = \dim H^m\mathscr{X}$$

du groupe $H^m\mathscr{X}$ (qui, rappelons-le, est un espace vectoriel) s'appelle *m-ième nombre de Betti* de la variété $\mathscr{X}$. (Ce nombre est souvent désigné par $b^m\mathscr{X}$.)

Il est évident que si $\dim \mathscr{X} = n$, alors

$$h^m\mathscr{X} = 0 \quad \text{pour } m > n$$

(puisque $\Omega^m\mathscr{X} = 0$ pour $m > n$). Donc, seuls les nombres de Betti

$$h^0\mathscr{X}, \ldots, h^n\mathscr{X}, \quad n = \dim \mathscr{X},$$

sont intéressants.

Signalons que *l'égalité $h^m\mathscr{X} = 0$ est nécessaire et suffisante pour que l'équation* (1) *admette une solution pour toute forme fermée* θ *de degré m sur la variété* $\mathscr{X}$.

* * *

Traitons d'abord le cas $m = 0$. (Ce cas mérite d'être étudié même si l'équation (1) n'a de sens que pour $m > 0$.)

Soit $\pi_0\mathscr{X}$ l'ensemble de toutes les composantes connexes de la variété différentiable $\mathscr{X}$.

Proposition 2. *L'espace vectoriel $H^0\mathscr{X}$ est naturellement isomorphe à l'espace vectoriel $\mathbb{R}^{\pi_0\mathscr{X}}$ de toutes les applications $\pi_0\mathscr{X} \to \mathbb{R}$. Donc, si l'ensemble $\pi_0\mathscr{X}$ est fini, le nombre zéro de Betti $h^0\mathscr{X}$ de $\mathscr{X}$ est égal au nombre $\#\pi_0\mathscr{X}$ de ses composantes connexes:*

$$h^0\mathscr{X} = \#\pi_0\mathscr{X}.$$

En particulier, $h^0\mathscr{X} > 0$ et $h^0\mathscr{X} = 1$ si et seulement si la variété $\mathscr{X}$ est connexe.

Prouvons tout d'abord un lemme géométrique.

Soit $\mathfrak{U} = \{U\}$ un recouvrement ouvert d'une variété $\mathscr{X}$. On dira que des éléments U et V de ce recouvrement sont enchaînés

si dans $\mathfrak{U}$ il existe une suite d'éléments $U_0, U_1, \ldots, U_m$, telle que $U_0 = U$, $U_m = V$, et pour tout $i = 1, \ldots, m$, les ensembles U_{i-1} et U_i se rencontrent:

$$U_{i-1} \cap U_i \neq \varnothing.$$

Lemme 1. *Toute composante connexe C de la variété $\mathscr{X}$ est contenue dans la réunion O d'une sous-famille du recouvrement $\mathfrak{U}$ composée de tous les éléments enchaînés avec l'un d'eux (donc enchaînés entre eux). Si tous les éléments du recouvrement $\mathfrak{U}$ sont connexes, alors $C = O$.*

Démonstration. Soient p_0 un point arbitraire de la composante C, et U_0 un élément du recouvrement $\mathfrak{U}$ tel que $p_0 \in U_0$. Supposons par ailleurs que O est la réunion de tous les éléments du recouvrement enchaînés avec U_0. Il est clair que O est ouvert (et contient le point p_0). Donc, pour prouver l'inclusion $C \subset O$ il suffit de montrer que O est aussi fermé.

Soit $p \in \bar{O}$ et soit U un élément du recouvrement $\mathfrak{U}$ tel que $p \in U$. Alors $U \cap O \neq \varnothing$ et, par suite, il existe un élément V de $\mathfrak{U}$ contenu dans O, tel que $U \cap V \neq \varnothing$. Donc, l'élément U est enchaîné à U_0 et, par conséquent, $U \subset O$. Donc, $p \in U \subset O$. Ce qui prouve que O est fermé.

Si tous les éléments du recouvrement $\mathfrak{U}$ sont connexes, il en est de même de O, puisque la réunion d'ensembles connexes sécants est connexe. Donc, $C = O$. □

Démonstration de la proposition 2. Par définition,

$$H^0\mathscr{X} = \operatorname{Ker}\{\mathrm{d} : \Omega^0\mathscr{X} \to \Omega^1\mathscr{X}\},$$

i.e.

$$H^0\mathscr{X} = \{f \in \mathsf{F}\mathscr{X}\,;\ \mathrm{d}f = 0\}.$$

Or la condition $\mathrm{d}f = 0$ exprime que dans toute carte $(U, h) = (U, x^1, \ldots, x^n)$ on a les égalités $\frac{\partial f}{\partial x^1} = 0, \ldots, \frac{\partial f}{\partial x^n} = 0$ et, par suite, si le voisinage de coordonnées U est connexe, alors $f = \text{const}$ sur U. Ceci étant, si $f = c_U$ sur U et $f = c_V$ sur V, on a nécessairement $c_U = c_V$ pour $U \cap V \neq \varnothing$.

Ayant ceci en vue, considérons un recouvrement $\mathfrak{U}$ de la variété $\mathscr{X}$ composé de tous les voisinages de coordonnées connexes. D'après ce qui a été prouvé, la fonction f prend la même valeur sur deux éléments quelconques sécants de $\mathfrak{U}$. Elle prend donc les mêmes valeurs sur deux éléments enchaînés quelconques de $\mathfrak{U}$. D'après le lemme 1 il s'ensuit que la fonction f est constante sur chaque composante connexe de la variété $\mathscr{X}$. Puisque inversement toute fonction f constante sur toute composante connexe de la variété $\mathscr{X}$ est différentiable sur $\mathscr{X}$ et satisfait à la relation $\mathrm{d}f = 0$, ceci prouve que

l'espace vectoriel $H^0\mathscr{X}$ est composé de toutes les fonctions $f: \mathscr{X} \to \mathbb{R}$ constantes sur chaque composante connexe de la variété $\mathscr{X}$ et, par suite, est naturellement isomorphe à l'espace vectoriel de toutes les applications $\pi_0\mathscr{X} \to \mathbb{R}$. □

* * *

Le calcul des nombres $h^m\mathscr{X}$ pour $m > 0$ est en principe un problème assez compliqué. Pour exposer les principes fondamentaux de sa résolution, il nous faut en premier lieu comparer les groupes de cohomologie de toute variété $\mathscr{X}$ avec ceux de la variété $\mathscr{X} \times \mathring{I}$, où $\mathring{I} = \mathbb{B}^1$ est l'intervalle $]-1, 1[$.

On rappelle (cf. définition 2 de la leçon 15) que les points de la variété $\mathscr{X} \times \mathring{I}$ sont les couples (p, t), où $p \in \mathscr{X}$, $-1 < t < 1$, et pour toute carte (U, h) de $\mathscr{X}$ le couple $(U \times \mathring{I}, h \times \mathrm{id})$, où

$$(h \times \mathrm{id})(p, t) = (h(p), t) \in \mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R},$$

est une carte de la variété $\mathscr{X} \times \mathring{I}$, et de plus toutes les cartes de cette forme constituent un atlas sur $\mathscr{X} \times \mathring{I}$. Si $x^1, \ldots, x^n$ sont des coordonnées locales de la carte (U, h), les coordonnées locales de la carte $(U \times \mathring{I}, h \times \mathrm{id})$ seront les fonctions $x^1 \circ \pi, \ldots, x^n \circ \pi, t$, où π est la projection $(p, t) \mapsto p$. Du reste, au lieu de $x^1 \circ \pi, \ldots, x^n \circ \pi$ on écrit généralement $x^1, \ldots, x^n$ sans crainte de confusion.

Si les coordonnées locales $x^1, \ldots, x^n$ et $x^{1'}, \ldots, x^{n'}$ des cartes (U, h) et (U', h') de la variété $\mathscr{X}$ sont reliées sur $U \cap U'$ par les relations

$$x^{i'} = x^{i'}(x^1, \ldots, x^n), \quad i' = 1, \ldots, n,$$

les coordonnées locales $x^1, \ldots, x^n, t$ et $x^{1'}, \ldots, x^{n'}, t$ des cartes $(U \times \mathring{I}, h \times \mathrm{id})$ et $(U' \times \mathring{I}, h' \times \mathrm{id})$ de la variété $\mathscr{X} \times \mathring{I}$ seront reliées sur $(U \cap U') \times \mathring{I} = (U \times \mathring{I}) \cap (U' \times \mathring{I})$ par les relations

$$\begin{aligned} x^{i'} &= x^{i'}(x^1, \ldots, x^n), \quad i' = 1, \ldots, n, \\ t &= t. \end{aligned} \tag{4}$$

On en déduit que pour toute carte (U, h) et tout point $q_0 = (p_0, t_0) \in U \times \mathring{I}$ le dernier vecteur $\left(\frac{\partial}{\partial t}\right)_{q_0}$ de la base

$$\left(\frac{\partial}{\partial x^1}\right)_{q_0}, \ldots, \left(\frac{\partial}{\partial x^n}\right)_{q_0}, \left(\frac{\partial}{\partial t}\right)_{q_0}$$

de l'espace $\mathrm{T}_{q_0}(\mathscr{X} \times \mathring{I})$ est *le même* pour toutes les cartes (U, h) avec $p_0 \in U$. Donc, la correspondance

$$q_0 \mapsto \left(\frac{\partial}{\partial t}\right)_{q_0}$$

définit de façon unique un champ de vecteurs $\frac{\partial}{\partial t}$ sur la variété $\mathscr{X} \times \mathring{I}$.

De façon analogue, le covecteur $(\mathrm{d}t)_{q_0}$ de la base duale

$$(\mathrm{d}x^1)_{q_0}, \ldots, (\mathrm{d}x^n)_{q_0}, (\mathrm{d}t)_{q_0}$$

de l'espace $\mathrm{T}^*_{q_0}(\mathscr{X} \times \mathring{I})$ ne dépend pas non plus du choix de la carte (U, h). Donc, la correspondance $q_0 \mapsto (\mathrm{d}t)_{q_0}$ définit de façon unique une forme différentielle linéaire $\mathrm{d}t$ sur $\mathscr{X} \times \mathring{I}$ (qui n'est autre que la différentielle de la fonction différentiable $(p, t) \mapsto t$).

Parmi les formes différentielles sur $\mathscr{X} \times \mathring{I}$ on distinguera les formes θ qui dans chaque carte $(U \times \mathring{I}, x^1, \ldots, x^n, t)$ se représentent par

$$\theta = \sum_{1 \leqslant i_1 < \ldots < i_m \leqslant n} \ldots \sum \theta_{i_1 \ldots i_m} \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_m}, \tag{5}$$

où $\theta_{i_1 \ldots i_m}$ sont des fonctions différentiables de $x^1, \ldots, x^n, t$. On dira de ces formes qu'elles sont *indépendantes de* $\mathrm{d}t$.

Pour tout t fixe, $-1 < t < 1$, la forme (5) peut être traitée comme une forme différentielle sur un voisinage de coordonnées U de la variété $\mathscr{X}$ et il est clair, étant donné l'identité des formules de changement des coordonnées $x^1, \ldots, x^n$ sur $\mathscr{X} \times \mathring{I}$ et $\mathscr{X}$, que ces formes (construites pour toutes les cartes (U, h)) sont compatibles sur les intersections, donc engendrent une forme θ_t sur $\mathscr{X}$.

Il est évident que ceci établit une correspondance biunivoque entre les formes θ sur $\mathscr{X} \times \mathring{I}$ indépendantes de $\mathrm{d}t$ et les familles $\{\theta_t\}$ de formes sur $\mathscr{X}$ dépendant *différentiablement* de t, $-1 < t < 1$ (i.e. telles que dans chaque voisinage de coordonnées leurs coefficients sont des fonctions différentiables de t). En principe, on identifiera θ et $\{\theta_t\}$.

Pour toute forme θ indépendante de $\mathrm{d}t$, et toutes les cartes de la forme $(U \times \mathring{I}, x^1, \ldots, x^n, t)$, les formes

$$\int_0^t \theta \, \mathrm{d}t = \sum_{1 \leqslant i_1 < \ldots < i_m \leqslant n} \ldots \sum \left(\int_0^t \theta_{i_1 \ldots i_m} \mathrm{d}t \right) \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_m}$$

sont compatibles sur les intersections et, par suite, engendrent une forme $\int_0^t \theta \, \mathrm{d}t$ indépendante aussi de $\mathrm{d}t$.

De façon analogue, on définit la forme $\frac{\partial \theta}{\partial t}$ dans chaque carte $(U \times \mathring{I}, x^1, \ldots, x^n, t)$ par la formule

$$\frac{\partial \theta}{\partial t} = \sum_{1 \leqslant i_1 < \ldots < i_m \leqslant n} \ldots \sum \frac{\partial \theta_{i_1 \ldots i_m}}{\partial t} \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_m}.$$

Ceci étant,

$$\frac{\partial}{\partial t}\left(\int_0^t \theta\, dt\right) = \theta$$

pour toute forme θ indépendante de dt.

De plus, la forme $d\theta_t$ est définie pour tout t fixe sur $\mathscr{X}$. Désignons par $\partial\theta$ la famille de formes $d\theta_t$, $-1 < t < 1$, traitée comme une forme sur $\mathscr{X} \times \mathring{I}$. La forme $\partial\theta$ est de degré $m + 1$, où m est le degré de la forme θ, et est reliée à la différentielle extérieure $d\theta$ de la forme θ par la formule

$$d\theta = \partial\theta + (-1)^m \frac{\partial\theta}{\partial t} \wedge dt. \tag{6}$$

En particulier,

$$d\int_0^t \theta\, dt = \partial \int_0^t \theta\, dt + (-1)^m \theta \wedge dt. \tag{7}$$

Parmi les formes différentielles sur $\mathscr{X} \times \mathring{I}$ indépendantes de dt, on distinguera les formes *indépendantes de t*, i.e. telles que dans leurs expressions (5) les coefficients $\theta_{i_1 \ldots i_m}$ soient indépendants de t. Ces formes s'identifient de façon naturelle aux formes sur $\mathscr{X}$ (en fait elles se représentent par $\pi^*\theta$, où θ est une forme sur $\mathscr{X}$, et π une projection de $\mathscr{X} \times \mathring{I}$ sur $\mathscr{X}$), de sorte qu'en vertu de cette identification,

$$\Omega^m \mathscr{X} \subset \Omega^m (\mathscr{X} \times \mathring{I}) \quad \text{pour tout } r \geqslant 0. \tag{8}$$

Les formes indépendantes de t sont caractérisées de toute évidence par la relation $\frac{\partial\theta}{\partial t} = 0$ et, par suite, sont telles que $d\theta = \partial\theta$. Ceci exprime que les injections (8) forment une application de cochaînes

$$\Omega^\cdot \mathscr{X} \to \Omega^\cdot (\mathscr{X} \times \mathring{I}) \tag{9}$$

et, par conséquent, pour tout $m \geqslant 0$ induisent l'homomorphisme

$$H^m \mathscr{X} \to H^m (\mathscr{X} \times \mathring{I}) \tag{10}$$

des groupes de cohomologie (qui n'est autre que l'homomorphisme π^* induit par la projection π).

Proposition 3. *L'homomorphisme* (10) *est un isomorphisme pour tout* $m \geqslant 0$.

D é m o n s t r a t i o n. Il est aisé de voir qu'une forme arbitraire ω sur $\mathscr{X} \times \mathring{I}$ se représente d'une seule façon par

$$\omega = \theta_1 \wedge dt + \theta_2, \tag{11}$$

où θ_1 et θ_2 sont des formes indépendantes de dt. (La représentation (11) est visiblement possible dans tout voisinage de coordonnées U et il est clair que les formes θ_1 et θ_2 obtenues sont compatibles sur les intersections.) En posant

$$\theta_3 = (-1)^m \int_0^t \theta_1 \, dt,$$

on trouve en vertu de la formule (7) que

$$\omega = d\theta_3 + \theta,$$

où

$$\theta = \theta_2 - \partial\theta_3$$

est une forme indépendante de dt.

Ceci prouve, lorsque la forme ω est fermée, que

$$[\omega] = [\theta],$$

i.e. toute classe de cohomologie de la variété $\mathscr{X} \times \mathring{I}$ contient une forme θ indépendante de dt. La forme θ est fermée comme ω, i.e.

$$\partial\theta + \frac{\partial\theta}{\partial t} \wedge dt = 0$$

et en particulier $\frac{\partial\theta}{\partial t} = 0$. Donc, la forme θ est en fait indépendante de t aussi, i.e. est l'image par π^* d'une forme sur $\mathscr{X}$. Ceci prouve que l'application (10) est un épimorphisme.

D'autre part, si une forme θ indépendante de t se représente par $d\omega$, où ω est une forme sur $\mathscr{X} \times \mathring{I}$, en représentant ω par la formule (11) et en remarquant que

$$d(\theta_1 \wedge dt + \theta_2) = \left(\partial\theta_1 + (-1)^m \frac{\partial\theta_2}{\partial t}\right) \wedge dt + \partial\theta_2,$$

on trouve aussitôt que

$$\partial\theta_1 + (-1)^m \left(\frac{\partial\theta_2}{\partial t}\right) = 0 \quad \text{et} \quad \partial\theta_2 = \theta.$$

L'égalité $\partial\theta_2 = \theta$ a lieu identiquement par rapport à t et en particulier pour $t = 0$. Puisque θ est indépendante de t, ceci prouve que $\partial(\theta_2)_0 = \theta$, où $(\theta_2)_0$ est la valeur de la forme θ_2 pour $t = 0$, et, par conséquent, $d(\theta_2)_0 = \theta$ dans $\mathscr{X}$. L'application (10) est donc un isomorphisme. □

Soit $\mathring{I}^n$ un cube ouvert de l'espace $\mathbb{R}^n$ composé des points $\mathbf{t} = (t_1, \ldots, t_n)$ tels que $-1 < t_1 < 1, \ldots, -1 < t_n < 1$.

Corollaire 1 (lemme de Poincaré). *Toute forme différentielle fermée ω de degré $m > 0$ sur le cube $\mathring{I}^n$ est une forme exacte, i.e.*

$$h^m \mathring{I}^n = 0 \quad \textit{pour } m > 0.$$

Démonstration. Il suffit de remarquer que

$$\mathring{I}^n = \{\mathrm{pt}\} \times \underbrace{\mathring{I} \times \ldots \times \mathring{I}}_{n \text{ fois}},$$

où $\{\mathrm{pt}\}$ est une variété 0-dimensionnelle composée d'un seul point, et d'appliquer n fois la proposition 3. (Il est clair que $H^m\{\mathrm{pt}\} = 0$ pour $m > 0$.) □

Exercice 1. Montrer que *le cube $\mathring{I}^n$ est difféomorphe à la boule ouverte* $\mathring{\mathbb{B}}^n = \{x \in \mathbb{R}^n;\ |x| > 1\}$. [Nota. Envisager la restriction à $\mathring{\mathbb{B}}^n$ de l'homéomorphisme $\mathbb{B}^n \to I^n$ construit dans la démonstration du corollaire 1 au théorème 2 de la leçon 9.]

Exercice 2. Montrer que *le cube $\mathring{I}^n$ est difféomorphe à l'espace* $\mathbb{R}^n$. [Nota. Construire un difféomorphisme $\mathbb{R}^n \to \mathring{\mathbb{B}}^n$.]

Il s'ensuit que dans le corollaire 1 le cube $\mathring{I}^n$ peut être remplacé aussi bien par la boule $\mathring{\mathbb{B}}^n$ que par l'espace $\mathbb{R}^n$.

* * *

Utilisons le lemme de Poincaré pour calculer les groupes de cohomologie $H^m S^n$ de la sphère S^m.

Supposons d'abord que $n = 2$ et $m = 1$.

On sait (cf. exemple 9 de la leçon 6 pour $n = 2$) que la sphère S^2 est munie d'un atlas composé de six cartes que nous désignerons par (U_k, h_k), où $k = \pm 1, \pm 2, \pm 3$. Les supports

$$U_{-3},\ U_{-2},\ U_{-1},\ U_1,\ U_2,\ U_3$$

de ces cartes sont des hémisphères composés des points $(x, y, z) \in S^2$ pour lesquels respectivement

$$x < 0, \quad y < 0, \quad z < 0, \quad z > 0, \quad y > 0, \quad x > 0.$$

Chacun de ces supports étant difféomorphe à un disque ouvert $\mathring{\mathbb{B}}^2$, le lemme de Poincaré nous dit que pour toute forme différentielle linéaire fermée α sur S^2, il existe des fonctions différentiables

$$f_k : U_k \to \mathbb{R}, \quad k = \pm 1, \pm 2, \pm 3,$$

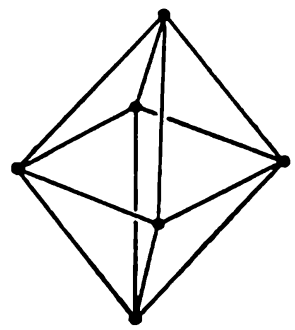
telles que $\alpha = df_k$ sur U_k. Ceci étant, pour tous k et l (pour lesquels l'intersection $U_k \cap U_l$ n'est pas vide) les formes df_k et df_l sont confondues sur $U_k \cap U_l$, donc, vu que l'ensemble $U_k \cap U_l$ est connexe, il existe une constante c_{kl} telle que

$$f_k - f_l = c_{kl} \quad \text{sur } U_k \cap U_l. \tag{12}$$

Si l'on convient de représenter les voisinages U_k par des points et que l'on joigne les points correspondant aux voisinages sécants par

des segments, le schéma combinatoire des intersections des voisinages se représente par un diagramme de la forme

(13)



qui n'est autre qu'un système d'arêtes d'un octaèdre. Les faces de cet octaèdre correspondent aux triplets de voisinages U_k dont l'intersection n'est pas vide.

Prouvons maintenant le lemme combinatoire suivant.

Lemme 2. *Supposons qu'à chaque arête orientée de l'octaèdre* (13) *est associé un nombre c_{kl} aux conditions suivantes*:

a) *aux arêtes kl et lk sont associés des nombres opposés*:

$$c_{lk} = -c_{kl}; \tag{14}$$

b) *pour toute face klm on a la relation*

$$c_{kl} + c_{lm} + c_{mk} = 0. \tag{15}$$

Il existe alors des nombres b_k tels que

$$b_l - b_k = c_{kl} \tag{16}$$

pour toute arête kl.

D é m o n s t r a t i o n. Il nous faut démontrer que le système de 12 équations linéaires (16) (par rapport aux 6 inconnues b_k) est compatible. Commençons tout d'abord par réduire le nombre de ces équations en utilisant les relations (15). Pour toute face *klm* de l'octaèdre (13), la relation (15) affirme que l'une des trois équations (16) correspondant aux arêtes de cette face est la conséquence des deux autres. Donc, chacune de ces relations permet de diminuer d'une unité le nombre des équations (16). En se servant de toutes ces relations (et en éliminant, pour fixer les idées, les équations correspondant aux arêtes aboutissant aux sommets 1 et —1), on obtient quatre équations

$$\begin{aligned} b_{-2} - b_{-3} &= c_{\bar{3}\bar{2}}, \quad & b_3 - b_{-2} &= c_{\bar{2}3}, \\ b_2 - b_3 &= c_{32}, \quad & b_{-3} - b_2 &= c_{2\bar{3}} \end{aligned} \tag{17}$$

en les quatre autres inconnues b_{-3}, b_{-2}, b_2, b_3 (ici et dans la suite on écrira parfois $\bar{1}$, $\bar{2}$, $\bar{3}$ pour —1, —2, —3). En posant par exemple

$b_{-3} = 0$, on obtient pour les inconnues b_{-2}, b_2 et b_3 les équations

$$b_{-2} = c_{\bar{3}\bar{2}}, \quad b_3 - b_{-2} = c_{\bar{2}3}, \quad b_2 - b_3 = c_{32},$$

d'où on les déduit immédiatement. [La raison de tout ceci c'est que les équations (17) sont linéairement dépendantes: leur somme est identiquement égale à zéro (vérifiez-le !). La cause géométrique de la simplification des inconnues est que les équations (17) correspondent aux arêtes de l'octaèdre (13) qui forment un chemin fermé, celle de la simplification des seconds membres est que ce chemin est le bord d'une pyramide composée de faces de sommet au point 1.]

Donc, si les conditions (15) (et 14)) sont remplies, les équations (16) sont bien compatibles. □

Remarque 1. Le lemme 2 reste de toute évidence en vigueur — et sa démonstration aussi — si c_{kl} sont des éléments d'un groupe (noté additivement), par exemple d'un espace vectoriel. On entend par là que b_k appartiennent à ce groupe (espace vectoriel).

Puisque les nombres c_{kl} définis par la formule (12) vérifient de toute évidence les relations (14) et (15), il existe en vertu du lemme 2 des nombres b_k, $k = \pm 1, \pm 2, \pm 3$, tels que pour toute arête kl de l'octaèdre (13)

$$f_k - f_l = b_l - b_k \quad \text{sur} \quad U_k \cap U_l,$$

i.e.

$$f_k + b_k = f_l + b_l \quad \text{sur} \quad U_k \cap U_l.$$

Il s'ensuit qu'en posant

$$f = f_k + b_k \quad \text{sur} \quad U_k$$

on définit de façon unique une fonction différentiable f sur $\mathcal{S}^2$. Ceci étant, $\mathrm{d}f = \mathrm{d}\,(f_k + b_k) = \mathrm{d}f_k = \alpha$ sur U_k et, par suite, $\mathrm{d}f = \alpha$ sur $\mathcal{S}^2$.

Ceci prouve que *toute forme différentielle fermée α de degré* 1 *sur $\mathcal{S}^2$ est une forme exacte*. Ce qui exprime que $H^1\mathcal{S}^2 = 0$, i.e.

$$h^1\mathcal{S}^2 = 0. \tag{18}$$

Remarquons qu'en plus du lemme de Poincaré on s'est servi du lemme combinatoire 2.

* * *

Essayons maintenant de calculer par la même méthode le groupe $H^1\mathcal{S}^1$, où $\mathcal{S}^1$ est le cercle $x^2 + y^2 = 1$.

Soient (voir p. 292)

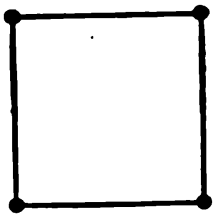
$$U_{-2}, U_{-1}, U_1, U_2$$

des demi-cercles définis respectivement par les inéquations

$$x < 0, \quad y < 0, \quad y > 0, \quad x > 0.$$

Le schéma des intersections des demi-cercles U_k, $k = \pm 1, \pm 2$, est un carré

(19)



D'après le lemme de Poincaré, pour toute forme linéaire ω sur S^1 (qui remarquons-le est nécessairement fermée) et tout $k = \pm 1, \pm 2$ il existe des fonctions f_k sur U_k, telles que

$$\omega = df_k \quad \text{sur} \quad U_k.$$

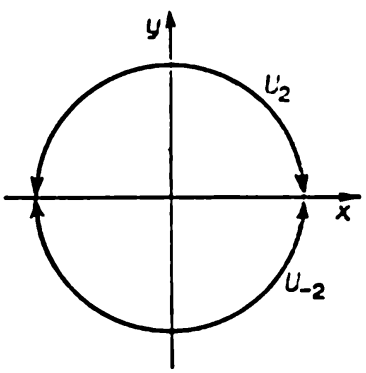
Ceci étant, pour tout côté kl du carré (19),

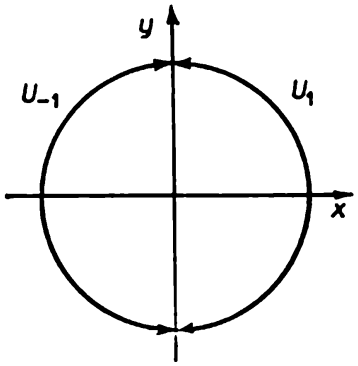
$$d(f_k - f_l) = 0 \quad \text{sur} \quad U_k \cap U_l,$$

et par suite, puisque l'ensemble $U_k \cap U_l$ est connexe,

$$f_k - f_l = c_{kl},$$

où c_{kl} sont des constantes. Ces constantes vérifient comme précédemment les relations (14) (les relations (15) n'ont pas de sens pour elles).





Les fonctions $c : kl \to c_{kl}$ définies sur les côtés du carré (19) et vérifiant la relation (14) (i.e. telles que $c_{lk} = -c_{kl}$) seront appelées *cocycles à une dimension du carré* (19). (De façon analogue, les fonctions $c : kl \mapsto c_{kl}$, définies sur les arêtes de l'octaèdre (13) et vérifiant les relations (14) et (15), s'appellent *cocycles à une dimension de l'octaèdre* (13); plus haut nous nous sommes passés de ce terme.) Il est clair que tous les cocycles $c : kl \mapsto c_{kl}$ forment un espace vectoriel Z^1 à 4 dimensions.

Les cocycles c pour lesquels il existe une application $b : k \mapsto b_k$ telle que

(20) $$b_k - b_l = c_{kl}$$

pour tout côté kl du carré (19) s'appellent *cobords*. Ils forment un sous-espace B^1 de l'espace Z^1. Désignons par H^1 l'espace quotient Z^1/B^1. Ses éléments (les classes de l'espace Z^1 suivant le sous-espace B^1) s'appellent *classes de cohomologie* du carré (19). Les cocycles appartenant à une même classe de cohomologie sont dits *cohomologues*.

[La terminologie nous suggère que nous avons en fait affaire à un complexe de cochaînes au sens de la définition 1. On développera cette idée à la leçon suivante.]

La construction d'un cocycle $c \in Z^1$ à l'aide d'une forme $\omega \in \Omega^1\mathcal{S}^1$ contient un élément d'arbitraire introduit par le choix des fonctions f_k. Mais si $\omega = df_k$ et $\omega = d\hat{f}_k$ sur U_k, alors $d(\hat{f}_k - f_k) = 0$ sur U_k et donc, $\hat{f}_k = f_k + b_k$, où b_k sont des constantes. Par conséquent, si $c_{kl} = f_k - f_l$ et $\hat{c}_{kl} = \hat{f}_k - \hat{f}_l$ sur $U_k \cap U_l$, alors

$$\hat{c}_{kl} - c_{kl} = b_k - b_l,$$

i.e. les cocycles $c: kl \mapsto c_{kl}$ et $\hat{c}: kl \mapsto \hat{c}_{kl}$ sont cohomologues. Ceci montre que la formule

$$\omega \mapsto [c],$$

où $[c]$ est la classe de cohomologie du cycle c, définit de façon unique une application

$$\Omega^1\mathcal{S}^1 \to H^1. \tag{21}$$

Si $\omega = df$, pour fonctions f_k nous pouvons prendre les restrictions $f|_{U_k}$ de la fonction f. Les nombres c_{kl} étant visiblement nuls pour un tel choix des fonctions f_k, *l'application* (21) *sera nulle sur les formes exactes*. Elle induit donc une application

$$H^1\mathcal{S}^1 \to H^1. \tag{22}$$

Il s'avère que *l'application* (22) *est un monomorphisme*, i.e. si le cocycle $c : kl \mapsto c_{kl}$ associé à la forme $\omega \in \Omega^1\mathcal{S}^1$ est un cobord, la forme ω est exacte. En effet, si $c_{kl} = f_k - f_l$ sur $U_k \cap U_l$ et si $c_{kl} = b_k - b_l$, alors $f_k - b_k = f_l - b_l$ sur $U_k \cap U_l$ et, par suite, la formule

$$f = f_k - b_k \quad \text{sur} \quad U_k$$

définit de façon unique une fonction f sur $\mathcal{S}^1$. Ceci étant, puisque

$$\omega = df_k = d(f_k - b_k) = df \text{ sur } U_k,$$

alors $\omega = df$ sur $\mathcal{S}^1$ et, par conséquent, la forme ω est exacte. □

Jusqu'ici nous avons intégralement suivi le calcul du groupe $H^1\mathcal{S}^2$. Pour poursuivre l'analogie nous devons maintenant prouver un lemme identique au lemme 2 (qui *mutatis mutandis* affirme de

toute évidence que $H^1 = 0$ dans le cas de l'octaèdre (13)). A cet effet il nous faudra analyser plus attentivement les égalités (20) (qui sont de même forme que les égalités (16) du lemme 2).

Ces égalités constituent un système de quatre équations

$$(23)\qquad \begin{aligned} b_1 - b_2 &= c_{1\,2}, \\ b_2 - b_{-1} &= c_{2\,\bar{1}}, \\ b_{-1} - b_{-2} &= c_{\bar{1}\,\bar{2}}, \\ b_{-2} - b_1 &= c_{\bar{2}\,1} \end{aligned}$$

à quatre inconnues b_{-2}, b_{-1}, b_1, b_2 (comme plus haut on remplacera l'indice $-k$ par $\bar{k}$), et affirmer que le cocycle c est un cobord revient à affirmer que ces équations sont compatibles.

En additionnant les équations (23) et en posant

$$\operatorname{Ind} c = c_{1\,2} + c_{2\,\bar{1}} + c_{\bar{1}\,\bar{2}} + c_{\bar{2}\,1},$$

on trouve immédiatement qu'une condition nécessaire de compatibilité de ces équations est $\operatorname{Ind} c = 0$.

La fonction Ind est une application linéaire de Z^1 dans $\mathbb{R}$, et la condition nécessaire $\operatorname{Ind} c = 0$ exprime que cette application est nulle sur le sous-espace B^1. Donc, la formule

$$\operatorname{Ind} [c] = \operatorname{Ind} c,$$

où $[c]$ est la classe de cohomologie du cocycle c, définit de façon unique une application

$$(24)\qquad \operatorname{Ind} : H^1 \to \mathbb{R}.$$

Les trois premières équations (23) nous donnent successivement

$$(25)\qquad \begin{aligned} b_2 &= b_1 - c_{1\,2}, \\ b_{-1} &= b_1 - c_{1\,2} - c_{2\,\bar{1}}, \\ b_{-2} &= b_1 - c_{1\,2} - c_{2\,\bar{1}} - c_{\bar{1}\,\bar{2}}. \end{aligned}$$

Donc, les formules (25) nous donnent la solution des trois premières équations pour tout b_1. Si $\operatorname{Ind} c = 0$, le nombre b_{-2} défini par la dernière formule (25) est, de toute évidence, solution de la quatrième équation (23). Donc, la condition $\operatorname{Ind} c = 0$ est aussi une condition suffisante de compatibilité des équations (23). S'agissant de l'application (24) ceci exprime qu'elle est un monomorphisme. Comme cette application est visiblement un épimorphisme, on en déduit que c'est un *isomorphisme*, donc

$$\dim H^1 = 1.$$

[Donc, si $H^1 = 0$ pour l'octaèdre, $H^1 \approx \mathbb{R}$ pour le carré.]

Contrairement au cas du groupe $H^1\mathbb{S}^2$, ce résultat ne permet pas encore de définir entièrement le groupe $H^1\mathbb{S}^1$; ce résul-

tat implique seulement que soit $H^1\mathbb{S}^1 = 0$, soit $H^1\mathbb{S}^1 \approx \mathbb{R}$. En fait, c'est le deuxième cas qui a lieu, i.e.

$$h^1\mathbb{S}^1 = 1. \tag{26}$$

Pour le prouver, il suffit d'exhiber une forme $\omega_0 \in \Omega^1\mathbb{S}^1$ pour laquelle le cocycle $c : kl \mapsto c_{kl}$ ne soit pas un cobord, i.e. soit tel que $\operatorname{Ind} c \neq 0$.

Du cours du secondaire on sait que la paramétrisation du cercle $\mathbb{S}^1$ consiste à associer à un nombre $t \in \mathbb{R}$ le point $p = (\cos t, \sin t)$ de ce cercle (géométriquement, la coordonnée t est l'angle formé par le rayon vecteur du point p avec l'axe positif des abscisses; pour cette raison on l'appelle *coordonnée angulaire* sur le cercle). La coordonnée t — dûment bornée — est une coordonnée locale au sens de la définition 1 de la leçon 6 sur chaque voisinage de coordonnées U_k (mais bien sûr pas sur le cercle $\mathbb{S}^1$ tout entier). Plus exactement, la coordonnée t bornée sur les intervalles

$$]0, \pi[, \ ]-\pi, 0[, \ \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[, \ \left]\frac{\pi}{2}, \frac{3\pi}{2}\right[$$

sera coordonnée locale respectivement sur

$$U_1, \ U_{-1}, \ U_2, \ U_{-2}.$$

Désignons cette coordonnée locale par t_k.

Donc,

$$\begin{aligned} &t_1 = t_{-2} \quad \text{sur } U_1 \cap U_{-2}, && t_{-1} = t_2 \quad \text{sur } U_{-1} \cap U_2, \\ &t_{-2} = t_{-1} + 2\pi \quad \text{sur } U_{-2} \cap U_{-1}, && t_2 = t_1 \quad \text{sur } U_2 \cap U_1. \end{aligned} \tag{27}$$

L'étude de la coordonnée t sur le cercle $\mathbb{S}^1$ tout entier exprime en fait que l'on passe à la droite $\mathbb{R}$ moyennant l'application

$$\mathbb{R} \to \mathbb{S}^1, \ t \mapsto (\cos t, \sin t).$$

Les fonctions sur $\mathbb{S}^1$ deviennent périodiques (de période 2π) sur $\mathbb{R}$, et les formes différentielles linéaires sur $\mathbb{S}^1$, des formes $f(t)\,dt$ sur $\mathbb{R}$ à coefficients périodiques $f(t)$.

Donc, en particulier, sur $\mathbb{S}^1$ est définie la forme $\omega_0 = dt$ (bien que t ne soit pas une fonction sur $\mathbb{S}^1$). Ceci étant,

$$dt = dt_k \quad \text{sur } U_k.$$

Donc, en vertu des relations (27), le cocycle $c = \{c_{kl}\}$ pour la forme ω_0 s'exprimera par les formules

$$c_{kl} = \begin{cases} 2\pi & \text{si } k = \bar{2}, \ l = \bar{1}, \\ -2\pi & \text{si } k = \bar{1}, \ l = \bar{2}, \\ 0 & \text{pour tous les autres } k \text{ et } l. \end{cases}$$

Donc, $\operatorname{Ind} c = 2\pi \neq 0$.

On peut construire la forme ω_0 sans recourir à la coordonnée angulaire t.

Il est clair que x et y sont des fonctions différentiables sur S^1, donc les formes différentielles dx et dy sont définies sur S^1.

Exercice 3. Montrer que

$$\omega_0 = x\,dy - y\,dx$$

et

$$\omega_0 = \begin{cases} -\dfrac{dx}{y} & \text{sur } U_1 \text{ et } U_{-1}, \\ \dfrac{dy}{x} & \text{sur } U_2 \text{ et } U_{-2}. \end{cases}$$

En déduire de nouveau que $\operatorname{Ind} c = 2\pi$.

* * *

Le groupe H^1S^1 peut être calculé par un procédé plus simple ne faisant pas intervenir les recouvrements $\{U_k\}$.

Ce procédé, qui fait appel à des raisonnements analytiques, consiste à associer à chaque forme $\omega = f(t)\,dt$ sur S^1 l'intégrale

$$I\omega = \int_0^{2\pi} f(t)\,dt.$$

L'application obtenue

$$\Omega^1S^1 \to \mathbb{R}, \quad \omega \mapsto I\omega, \tag{28}$$

est visiblement linéaire et épimorphe (car $I\omega_0 \neq 0$ pour la forme $\omega = dt$). Si $\omega = dg$, où g est une fonction sur S^1 (une fonction périodique sur $\mathbb{R}$), i.e. si $f(t) = g'(t)$, alors

$$I\omega = g(t)\,|_0^{2\pi} = 0.$$

Réciproquement, si $I\omega = 0$, la fonction

$$g(t) = \int_0^t f(t)\,dt$$

est périodique et, par conséquent, est une fonction sur S^1 telle que $dg = \omega$.

Donc, le noyau de l'application (28) est le sous-espace des formes exactes, et par conséquent, cette application induit un isomorphisme de H^1S^1 sur $\mathbb{R}$.

Donc, $h^1S^1 = 1$.

La concision et la simplicité de cette démonstration illustrent bien la puissance des méthodes du calcul intégral.

* * *

Calculons maintenant le groupe $H^2\mathbb{S}^2$.

Considérons de nouveau le recouvrement $\{U_k,\ k = \pm 1, \pm 2, \pm 3\}$ de la sphère $\mathbb{S}^2$, dont le schéma combinatoire des intersections est l'octaèdre (13), et soit ω une forme (automatiquement fermée) de degré 2 sur la sphère $\mathbb{S}^2$.

Chaque ensemble U_k étant difféomorphe à un disque $\mathring{\mathbb{B}}^2$, le lemme de Poincaré nous dit que pour tout $k = \pm 1, \pm 2, \pm 3$ il existe sur U_k une forme α_k de degré 1 telle que $\omega = d\alpha_k$ sur U_k. Pour tous k et l (tels que $U_k \cap U_l \neq \varnothing$) la forme $\alpha_k - \alpha_l$ est fermée sur $U_k \cap U_l$, et puisque l'ensemble $U_k \cap U_l$ est aussi difféomorphe au disque $\mathring{\mathbb{B}}^2$, le lemme de Poincaré affirme qu'il existe une fonction différentiable $f_{kl}: U_k \cap U_l \to \mathbb{R}$ telle que

$$\alpha_k - \alpha_l = df_{kl} \quad \text{sur} \quad U_k \cap U_l \tag{29}$$

(ceci étant, on peut, sans perdre en généralité, admettre de toute évidence que $f_{kl} = -f_{lk}$ pour tous k et l).

Pour tout triplet d'indices k, l, m (tels que $U_k \cap U_l \cap U_m \neq \varnothing$) on a l'égalité

$$d\,(f_{kl} + f_{lm} + f_{mk}) = \alpha_k - \alpha_l + \alpha_l - \alpha_m + \alpha_m - \alpha_k = 0,$$

donc (puisque l'ensemble $U_k \cap U_l \cap U_m$ est connexe) il existe des constantes c_{klm} telles que

$$f_{kl} + f_{lm} + f_{mk} = c_{klm} \quad \text{sur } U_k \cap U_l \cap U_m$$

(on peut, sans perdre en généralité, admettre de toute évidence que les nombres c_{klm} dépendent antisymétriquement des indices k, l et m, i.e. changent de signe par toute transposition de ces indices).

Cette construction recèle bien sûr un certain élément d'arbitraire. En effet, *primo*, au lieu des formes α_k on peut prendre des formes quelconques $\hat{\alpha}_k$ telles que $d\hat{\alpha}_k = d\alpha_k$, i.e. telles que (on se sert encore du lemme de Poincaré) $\hat{\alpha}_k = \alpha_k + dg_k$, où $g_k: U_k \to \mathbb{R}$ sont des fonctions différentiables. *Secundo*, les fonctions f_{kl} peuvent de façon analogue être remplacées par les fonctions $f_{kl} + b_{kl}$, où b_{kl} sont des constantes arbitraires (vérifiant la relation $b_{lk} = -b_{kl}$). La prise en compte de ces deux possibilités conduit à remplacer les fonctions f_{kl} par les fonctions $f_{kl} + g_k - g_l + b_{kl}$, et donc, les nombres c_{klm} par les nombres

$$\hat{c}_{klm} = c_{klm} + (b_{kl} + b_{lm} + b_{mk}). \tag{30}$$

(Remarquons que la dépendance par rapport aux fonctions g_k disparaît à la dernière étape.)

La forme ω définit donc une fonction antisymétrique $c: klm \mapsto c_{klm}$ à une équivalence près décrite par la relation (30).

On appellera les fonctions antisymétriques $c: klm \mapsto c_{klm}$ *cocycles à deux dimensions* de l'octaèdre (13), et les cocycles de la forme

$$c_{klm} = b_{kl} + b_{lm} + b_{mk}, \tag{31}$$

cobords. Les cocycles forment un espace vectoriel Z^2 (de dimension 8, soit le nombre des faces de l'octaèdre (13)), et les cobords, un sous-espace B^2 de Z^2. Désignons par H^2 l'espace quotient Z^2/B^2. Ses éléments, i.e. les classes de Z^2 suivant B^2, s'appellent *classes de cohomologie à deux dimensions* de l'octaèdre (13).

Puisque la relation (30) exprime que les cocycles $c: klm \mapsto c_{klm}$ et $\hat{c}: klm \mapsto \hat{c}_{klm}$ appartiennent à la même classe de cohomologie $[c]$, on voit qu'on a construit une application (visiblement linéaire)

$$\Omega^2\mathcal{S}^2 \to H^2, \quad \omega \mapsto [c]. \tag{32}$$

Si la forme ω est exacte, i.e. se représente par $d\alpha$, où $\alpha \in \Omega^1\mathcal{S}^2$, pour formes α_k on peut prendre les restrictions $\alpha|_{U_k}$ de la forme α aux voisinages de coordonnées U_k. Les différences $\alpha_k - \alpha_l$ seront nulles, et les relations (29) seront satisfaites si pour f_{kl} on prend des fonctions identiquement nulles. Les nombres c_{klm} étant visiblement nuls pour de telles fonctions f_{kl}, ceci prouve que *toutes les formes exactes se transforment en le zéro de l'espace H^2 par l'application* (32) (la classe de cohomologie $[0]$ du cocycle zéro), donc cette application induit une application (linéaire aussi)

$$H^2\mathcal{S}^2 \to H^2. \tag{33}$$

Par analogie avec le cas du groupe $H^1\mathcal{S}^1$, il est naturel de s'attendre à ce que *cette application soit un monomorphisme*, i.e. si le cocycle $c: klm \mapsto c_{klm}$ associé à la forme ω se représente par (31), la forme ω est exacte. Mais si la relation (31) est réalisée, en remplaçant les fonctions f_{kl} par les différences $f_{kl} - b_{kl}$, on trouve que $c_{klm} = 0$. Dans ce cas donc on peut, sans perdre en généralité, admettre que pour toute face klm de l'octaèdre (13) on a la relation

$$f_{kl} + f_{lm} + f_{mk} = 0 \quad \text{sur} \quad U_k \cap U_l \cap U_m. \tag{34}$$

Lemme 3. *Si les fonctions $f_{kl} = -f_{lk}$ satisfont à la relation* (34), *il existe des fonctions $g_k: U_k \to \mathbb{R}$ telles que*

$$f_{kl} = g_k - g_l \quad \textit{sur} \quad U_k \cap U_l$$

pour toute arête kl de l'octaèdre (13).

A la leçon 22 on prouvera un théorème général 1 dont un cas très spécial est le lemme 3. Pour l'instant nous adopterons ce lemme sans le prouver.

Exercice 4. Utiliser l'identité de forme de la relation (34) et de la relation (15) du lemme 2 et la remarque 1 pour prouver le lemme 3.

Du lemme 3 il s'ensuit que les formes $\beta_k = \alpha_k - dg_k$ vérifient pour toute arête kl de l'octaèdre (13) les relations

$$\beta_k = \beta_l \quad \text{sur } U_k \cap U_l$$

et engendrent donc une forme β sur $\mathcal{S}^2$. De plus, vu que pour tout k

$$d\beta_k = d\alpha_k = \omega \quad \text{sur} \quad U_k,$$

on en déduit que $d\beta = \omega$ sur $\mathcal{S}^2$, donc la forme ω et exacte.

Par conséquent, l'application (33) est bien un monomorphisme. □

La suite du calcul est connue. On montre tout d'abord que $H^2 \approx \approx \mathbb{R}$ (à l'aide de la fonctionnelle Ind: $Z^2 \mapsto \mathbb{R}$ qui est égale à la somme des nombres c_{klm} sur toutes les faces — dûment orientées! — de l'octaèdre (13)), puis on exhibe une forme ω_0 pour laquelle le cocycle c est tel que Ind $c \neq 0$. On trouve en définitive que $H^2\mathcal{S}^2 \approx \approx \mathbb{R}$, i. e.

$$h^2\mathcal{S}^2 = 1. \tag{35}$$

Exercice 5. Effectuer en détail la démonstration envisagée de l'égalité (35). [N o t a. Pour forme ω_0, prendre la forme

$$\omega_0 = x\,dy \wedge dz - y\,dx \wedge dz + z\,dx \wedge dy.$$

La façon la plus simple pour démontrer que cette forme n'est pas exacte est d'abord de prouver par la formule de Stokes (ou, à la rigueur, par la formule de Green) que *l'intégrale d'une forme exacte étendue à la sphère* $\mathcal{S}^2$ *est nulle*, et ensuite d'établir que l'intégrale de la forme ω_0 est égale à 4π.]

* * *

On voit que même pour les variétés les plus simples le calcul des groupes de cohomologie se heurte à des difficultés de nature aussi bien combinatoire qu'analytique. S'agissant des difficultés combinatoires, on peut parfois les surmonter — ou à la rigueur les alléger — par un choix convenable du recouvrement. Par exemple, le calcul du groupe $H^1\mathcal{S}^2$ se simplifie singulièrement si l'on se sert de l'atlas à deux éléments $\{(U, h), (V, k)\}$ de l'exemple 10 de la leçon 6. Au lieu de l'octaèdre on obtiendrait le segment

et la démonstration de l'égalité (18) serait triviale. [D'après le lemme de Poincaré, $\alpha = df_U$ sur U, et $\alpha = df_V$ sur V, et de plus $f_U - - f_V = b$ sur $U \cap V$, où b est une constante. Donc, la formule

$$f = \begin{cases} f_U & \text{sur } U, \\ f_V + b & \text{sur } V \end{cases}$$

définit de façon unique une fonction f sur $\mathcal{S}^2$, telle que $df = \alpha$.]

Puisque la dernière démonstration s'étend *ad litteram* au cas d'une sphère S^n de dimension quelconque $n \geqslant 2$, on trouve même que

$$h^1 S^n = 0 \quad \text{pour } n \geqslant 2. \tag{36}$$

Mais si l'on tente par cette méthode de calculer les nombres $h^m S^n$ pour $m \geqslant 2$, on se heurte entre autres difficultés à celle que le lemme de Poincaré n'est pas directement applicable sur l'intersection $U \cap V$ des cartes U et V (puisque cette intersection est difféomorphe à l'espace pointé $\mathbb{R}^n \setminus \{0\}$ et pas à une boule). N'empêche que ce calcul est possible si à la place du lemme de Poincaré on se sert de la proposition 3 et que l'on raisonne par récurrence sur n.

* * *

Traitons d'abord le cas où $0 < m < n$.

Proposition 4. *Pour $0 < m < n$ on a l'égalité*

$$h^m S^n = 0. \tag{37}$$

D é m o n s t r a t i o n. Pour $n = 2$ (m est alors nécessairement égal à 1) l'égalité (37) a déjà été prouvée (cf. (18) ou (36)). Effectuons une récurrence sur n. On peut admettre que $m \geqslant 2$, car pour $m = 1$ l'égalité (37) a déjà été prouvée plus haut.

Supposons qu'on ait déjà prouvé que $h^m S^{n-1} = 0$ pour $0 < m < n - 1$ et soit ω une forme fermée sur S^n de degré m, où $0 < m < n$. Considérons le recouvrement à deux éléments $\{U, V\}$ de la sphère S^n de l'exemple 10 de la leçon 6.

Exercice 6. Montrer que *l'ensemble $U \cap V = S^n \setminus \{p_0, q_0\}$ est difféomorphe au produit $S^{n-1} \times \mathring{I}$.* [N o t a. Construire les difféomorphismes $U \cap V \to \mathbb{R}^n \setminus \{0\}$ et $S^{n-1} \times \mathring{I} \to \mathbb{R}^n \setminus \{0\}$.]

En vertu du lemme de Poincaré, il existe sur U et V des formes θ_U et θ_V de degré $m - 1$, telles que $\omega = d\theta_U$ sur U et $\omega = d\theta_V$ sur V. La forme $\theta_U - \theta_V$ est fermée sur l'intersection $U \cap V$ et est de degré $m - 1$. Mais d'après l'assertion de l'exercice 6 et la proposition 3, le nombre de Betti $h^{m-1} S^{n-1}$ de l'intersection $U \cap V$ est égal au nombre de Betti $h^{m-1} S^{n-1}$ de la sphère S^{n-1}, donc à zéro d'après l'hypothèse de la récurrence. Il existe par conséquent sur $U \cap V$ une forme α de degré $m - 2$, telle que $\theta_U - \theta_V = d\alpha$.

Exercice 7. Construire sur la sphère S^n une fonction différentiable f prenant ses valeurs dans l'intervalle $[0, 1]$ et égale à 0 à proximité d'un point q_0 (i.e. dans un voisinage de ce point) et à 1 à proximité d'un point p_0. [N o t a. Cf. le corollaire 2 au lemme 1 de la leçon 1 ou la proposition 2 de la leçon 14.]

Définissons à l'aide de la fonction f de l'exercice 7 une forme α_U sur U et une forme α_V sur V en posant

$$(\alpha_U)_p = \begin{cases} f(p)\alpha_p & \text{si } p \in U \cap V, \\ 0 & \text{sinon (i.e. si } p = p_0), \end{cases}$$

$$(\alpha_V)_p = \begin{cases} (f(p)-1)\alpha_p & \text{si } p \in U \cap V, \\ 0 & \text{sinon (i.e. si } p = q_0). \end{cases}$$

Il est clair que les formes α_U et α_V sont différentiables et $\alpha = \alpha_U - \alpha_V$ sur $U \cap V$, donc $\theta_U - d\alpha_U = \theta_V - d\alpha_V$ sur $U \cap V$. Donc, la formule

$$\theta = \begin{cases} \theta_U - d\alpha_U & \text{sur } U, \\ \theta_V - d\alpha_V & \text{sur } V \end{cases}$$

définit de façon unique une forme θ sur S^n, telle que $d\theta = \omega$. La forme ω étant une forme fermée arbitraire de degré m sur S^n, ceci prouve que $h^m S^n = 0$.

* * *

Le cas $m = n$ se traite de façon analogue.

Proposition 5. *Pour tout $n \geqslant 1$ on a l'égalité*

$$h^n S^n = 1. \tag{38}$$

D é m o n s t r a t i o n. L'égalité (38) ayant déjà été prouvée pour $n = 1$ (cf. formule (26)), on peut de nouveau raisonner par récurrence sur n.

Soit $\{U, V\}$ le même recouvrement de la sphère S^n que plus haut. Puisque $h^{n-1}S^{n-1} = 1$ par hypothèse de la récurrence et donc que

$$h^{n-1}(U \cap V) = h^{n-1}(S^{n-1} \times \mathring{I}) = h^{n-1}S^{n-1} = 1,$$

il existe sur $U \cap V$ une forme fermée mais non exacte $\theta^{(0)}$ de degré $n-1$, telle que toute forme de degré $n-1$ sur $U \cap V$ est cohomologue à une forme $a\theta^{(0)}$, où a est un nombre.

Par la méthode déjà connue, on peut, en se servant de la fonction f de l'exercice 7, représenter la forme $\theta^{(0)}$ par

$$\theta^{(0)} = \theta_U^{(0)} - \theta_V^{(0)} \quad \text{sur } U \cap V, \tag{39}$$

où $\theta_U^{(0)}$ et $\theta_V^{(0)}$ sont des formes de degré $n-1$ sur U et V respectivement. Etant donné que par hypothèse $d\theta^{(0)} = 0$, il vient en vertu de la formule (39)

$$d\theta_U^{(0)} = d\theta_V^{(0)} \quad \text{sur } U \cap V,$$

donc les égalités

$$\omega^{(0)} = \begin{cases} d\theta_U^{(0)} & \text{sur } U, \\ d\theta_V^{(0)} & \text{sur } V \end{cases}$$

définissent de façon unique une forme $\omega^{(0)}$ de degré n sur S^n.

Supposons maintenant que ω est une forme arbitraire de degré n (nécessairement fermée) sur la sphère S^n. En vertu du lemme de Poincaré, il existe sur U et V des formes θ_U et θ_V de degré $n-1$, telles que $\omega = d\theta_U$ sur U et $\omega = d\theta_V$ sur V. La forme $\theta_U - \theta_V$ est fermée sur $U \cap V$, donc il existe sur $U \cap V$ une forme α de degré $n-2$ et un nombre $a \in \mathbb{R}$, tels que

$$\theta_U - \theta_V = d\alpha + a\theta^{(0)} \quad \text{sur} \quad U \cap V.$$

En représentant comme plus haut la forme α par $\alpha = \alpha_U - \alpha_V$, où α_U et α_V sont des formes sur U et V respectivement, on peut mettre cette égalité sous la forme suivante:

$$\theta_U - d\alpha_U - a\theta_U^{(0)} = \theta_V - d\alpha_V - a\theta_V^{(0)} \quad \text{sur } U \cap V.$$

Donc, la formule

$$\beta = \begin{cases} \theta_U - d\alpha_U - a\theta_U^{(0)} & \text{sur } U, \\ \theta_V - d\alpha_V - a\theta_V^{(0)} & \text{sur } V \end{cases}$$

définit de façon unique une forme β de degré $n-1$ sur S^n. Ceci étant,

$$d\beta = d\theta_U - a\, d\theta_U^{(0)} = \omega - a\omega^{(0)} \quad \text{sur } U$$

et

$$d\beta = d\theta_V - a\, d\theta_V^{(0)} = \omega - a\omega^{(0)} \quad \text{sur } V.$$

Donc, $d\beta = \omega - a\omega^{(0)}$ sur toute la sphère S^n, i.e. la forme ω est cohomologue à la forme $a\omega^{(0)}$.

Pour achever la démonstration de l'égalité (38) il reste donc à prouver seulement que la forme $\omega^{(0)}$ n'est pas cohomologue à 0 (i.e. n'est pas une forme exacte). Mais si $\omega^{(0)} = d\theta$, on a $d\theta_U^{(0)} = d(\theta|_U)$ sur U, donc $\theta_U^{(0)} = \theta|_U + d\varphi_U$, où φ_U est une forme de degré $n-2$ sur U. De façon analogue, $\theta_V^{(0)} = \theta|_V + d\varphi_V$, où φ_V est une fonction différentiable sur V.

Donc,

$$\theta^{(0)} = \theta_U^{(0)} - \theta_V^{(0)} = d(\varphi_U - \varphi_V) \quad \text{sur } U \cap V,$$

i.e. la forme $\theta^{(0)}$ est exacte, ce qui est absurde. Donc, l'égalité $\omega^{(0)} = d\theta$ est impossible. □

Sur l'exemple des sphères S^1, S^2 et S^n nous avons illustré trois méthodes essentielles de calcul des groupes de cohomologie d'une variété (à l'aide des recouvrements dont l'intersection de toutes sous-familles est difféomorphe à une boule, à l'aide de recouvrements plus généraux ne satisfaisant pas à cette condition et enfin à l'aide d'intégrales). Ces méthodes seront étudiées plus systématiquement dans les leçons suivantes. L'exemple des sphères nous servira de modèle et de point de départ même si en principe on n'y fera pas explicitement référence.

LEÇON 21

Schémas simpliciaux et leurs réalisations géométriques. — Groupes de cohomologie de schémas simpliciaux. — Complexe double d'un recouvrement. — Groupes de cohomologie d'un complexe double. — Complexes doubles bordés. — Homomorphismes bords. — Complexes acycliques. — Acyclicité par rapport aux lignes pour $p = 0$.

Tout d'abord nous allons décrire sous une forme abstraite les schémas combinatoires des intersections mutuelles des éléments des recouvrements de variétés différentiables (ou, de façon plus générale, d'espaces topologiques).

Soient $\mathscr{X}$ un espace topologique arbitraire et $\mathfrak{U} = \{U_k, k \in K\}$ un recouvrement ouvert de $\mathscr{X}$. On dit qu'un sous-ensemble fini $\{k_0, \ldots, k_m\}$ d'un ensemble d'indices K est *distingué* si l'intersection $U_{k_0} \cap \ldots \cap U_{k_m}$ n'est pas vide:

$$U_{k_0} \cap \ldots \cap U_{k_m} \neq \varnothing.$$

Définition 1. Un ensemble K dont certains sous-ensembles finis sont distingués s'appelle *schéma simplicial* si toute partie d'un sous-ensemble distingué est distinguée.

Il est clair que les sous-ensembles distingués de l'ensemble des indices K du recouvrement $\mathfrak{U}$ satisfont à cette condition et donc K est un schéma simplicial. Ce schéma s'appelle *nerf* du recouvrement $\mathfrak{U}$.

On dit que $m + 1$ points d'un espace affine sont *affinement indépendants* s'ils n'appartiennent à aucun plan à $m - 1$ dimensions, i.e. si les différences $\mathbf{k}_1 - \mathbf{k}_0, \ldots, \mathbf{k}_m - \mathbf{k}_0$ de leurs rayons vecteurs $\mathbf{k}_0, \mathbf{k}_1, \ldots, \mathbf{k}_m$ sont linéairement indépendantes.

Pour tous $m + 1$ points $\mathbf{k}_0, \ldots, \mathbf{k}_m$ affinement indépendants de l'espace $\mathbb{R}^n$ (ou, de façon plus générale, d'un espace vectoriel sur le corps $\mathbb{R}$), l'ensemble de tous les points $\mathbf{k}$ de la forme

$$\mathbf{k} = t_0\mathbf{k}_0 + \ldots + t_m\mathbf{k}_m,$$

où $t_0, \ldots, t_m$ sont des nombres réels tels que

$$0 \leqslant t_0 \leqslant 1, \ldots, 0 \leqslant t_m \leqslant 1$$

et

$$t_0 + \ldots + t_m = 1$$

(nous identifions comme toujours les points de $\mathbb{R}^n$ à leurs rayons vecteurs), s'appelle *simplexe à m dimensions de sommets* $\mathbf{k}_0, \ldots, \mathbf{k}_m$ et se désigne par le symbole $\mathbf{k}_0, \ldots, \mathbf{k}_m$. Pour $m=0$ c'est le point $\mathbf{k}_0$; pour $m = 1$, le segment $\mathbf{k}_0\mathbf{k}_1$; pour $m = 2$, le triangle $\mathbf{k}_0\mathbf{k}_1\mathbf{k}_2$; pour $m = 3$, le tétraèdre $\mathbf{k}_0\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3$.

Les nombres $t_0, \ldots, t_m$ s'appellent *coordonnées barycentriques* du point $\mathbf{k}$ du simplexe $\mathbf{k}_0 \ldots \mathbf{k}_m$.

Un point $\mathbf{k}$ est un *point intérieur* à un simplexe $\mathbf{k}_0 \ldots \mathbf{k}_m$ si $0 < t_i < 1$ pour tout $i = 0, \ldots, m$. (Un tel point est intérieur au simplexe $\mathbf{k}_0 \ldots \mathbf{k}_m$ au sens topologique général par rapport au plan à m dimensions contenant le simplexe.)

Un schéma simplicial K sera dit *réalisable* si ses éléments peuvent être représentés (pour n assez grand) par des points de l'espace $\mathbb{R}^n$, tels que :

1) pour tout ensemble distingué $\{k_0, \ldots, k_m\} \subset K$, les points correspondants $\mathbf{k}_0, \ldots, \mathbf{k}_m$ sont affinement indépendants (et définissent donc un simplexe $\mathbf{k}_0 \ldots \mathbf{k}_m$) ;

2) les simplexes $\mathbf{k}_0 \ldots \mathbf{k}_m$ et $\mathbf{k}'_0 \ldots \mathbf{k}'_m$ correspondant à deux sous-ensembles distingués distincts $\{k_0, \ldots, k_m\}$ et $\{k'_0, \ldots, k'_m\}$ du schéma K ne possèdent pas de points intérieurs communs.

Dans ce cas, la réunion de tous les simplexes $\mathbf{k}_0 \ldots \mathbf{k}_m$ correspondant à tous les sous-ensembles distingués $\{k_0, \ldots, k_m\}$ s'appelle *réalisation géométrique* du schéma K et se désigne par le symbole $|K|$.

Les figures tracées à la leçon précédente représentaient justement la réalisation géométrique des nerfs des recouvrements correspondants de la sphère S^2 et du cercle S^1.

Par une tendance au transfert des termes, qui est propre aux mathématiques, les sous-ensembles distingués du schéma K sont généralement appelés *simplexes* de K. [Pour faire la distinction — si besoin est — entre les simplexes dans K et les simplexes dans $\mathbb{R}^n$, on dit des premiers qu'ils sont *abstraits* (ou *schématiques*) et des seconds, qu'ils sont *euclidiens* (ou *géométriques*).] Les éléments d'un schéma figurant dans l'un au moins de ses simplexes s'appellent ses *sommets* et le simplexe $\{k_0, \ldots, k_m\}$, *simplexe de sommets* $k_0, \ldots, k_m$.

Les éléments du schéma K, qui ne sont pas des sommets, ne participeront dans la suite à aucune construction (tout comme ils ne participent à aucune réalisation géométrique $|K|$) et on peut les négliger sans crainte. [Pour cette raison, en définissant un schéma de nombreux auteurs ont introduit un axiome auxiliaire impliquant que chaque élément du schéma soit un sommet. Mais en définissant le nerf d'un recouvrement, il faut alors exiger que ce recouvrement soit composé d'ensembles non vides, ce qui n'est pas toujours commode.]

Remarque 1. En fait nous n'aurons nulle part besoin des réalisations géométriques. Nous les avons décrites uniquement pour pouvoir appliquer le langage géométrique aux schémas simpliciaux. C'est

pourquoi en particulier nous glisserons sur la question assez délicate de savoir quels schémas simpliciaux sont réalisables. Remarquons seulement que tout schéma simplicial fini est visiblement réalisable.

Exercice 1. On dira qu'un schéma simplicial K est *injectivement réalisable* s'il admet une réalisation géométrique $|K|$ telle qu'un sous-ensemble $F \subset |K|$ est fermé (pour la topologie induite) si et seulement si pour tout simplexe $\{k_0, \ldots, k_m\} \subset K$ est fermée l'intersection $F \cap \mathbf{k}_0 \ldots \mathbf{k}_m$. Montrer qu'un schéma simplicial est injectivement réalisable si et seulement si

a) l'ensemble de ses sommets est fini ou dénombrable;

b) chacun de ses sommets n'appartient qu'à un nombre fini de simplexes;

c) il existe un n tel que chaque simplexe du schéma K contient $n+1$ sommets au plus.

Remarque 2. Les sous-ensembles de l'espace $\mathbb{R}^n$, qui sont les réalisations géométriques de schémas simpliciaux, s'appellent *polyèdres*. Leur théorie topologique qui est appelée généralement t o p o - l o g i e l i n é a i r e p a r m o r c e a u x est étroitement liée à la théorie topologique des variétés différentiables (on démontre — un difficile théorème! — que toute variété séparée différentiable à base dénombrable est homéomorphe à un polyèdre) et constitue « la plus géométrique » des parties de la topologie moderne. Malheureusement, cette théorie sort totalement du cadre de ce cours.

* * *

Soit K un schéma simplicial et soit K_m, $m \geqslant 0$, un sous-ensemble du produit

$$\underbrace{K \times \ldots \times K}_{m+1 \text{ fois}},$$

composé des suites $(k_0, \ldots, k_m)$, $k_0, \ldots, k_m \in K$, telles que l'ensemble $\{k_0, \ldots, k_m\}$ est un simplexe du schéma K. (Les éléments de K_m s'appellent généralement *simplexes ordonnés à m dimensions* du schéma K.)

Soit par ailleurs G un groupe abélien (noté additivement).

Définition 2. L'application

$$c\colon K_m \to G, \quad (k_0, \ldots, k_m) \mapsto c(k_0, \ldots, k_m)$$

s'appelle *cochaîne à m dimensions du schéma K sur le groupe G* (où *à valeurs* dans le groupe G) si cette application est antisymétrique, i.e. si $c(k_0, \ldots, k_m)$ change de signe par permutation de deux arguments voisins k_a, k_{a+1}, $0 \leqslant a \leqslant m-1$. En d'autres termes, l'application c est une cochaîne si

$$c(k_{\sigma(0)}, \ldots, k_{\sigma(m)}) = \varepsilon_\sigma c(k_0, \ldots, k_m) \tag{1}$$

pour tout simplexe ordonné $(k_0, \ldots, k_m) \in K_m$ et toute permutation σ des indices $0, 1, \ldots, m$ (où comme toujours ε_σ est la signature de la permutation σ).

Ces cochaînes forment de toute évidence un groupe qui est généralement désigné par $C^m(K; G)$.

Lorsque G est un corps, le groupe $C^m(K; G)$ est un espace vectoriel sur ce corps. En particulier, pour $G = \mathbb{R}$ (le seul cas intéressant !) le groupe $C^m(K; \mathbb{R})$ est un espace vectoriel sur le corps $\mathbb{R}$.

Il est aisé de voir que pour toute cochaîne $c \in C^m(K; G)$ la formule

$$(2) \qquad (\delta c)(k_0, \ldots, k_{m+1}) = \sum_{i=0}^{m+1} (-1)^i c(k_0, \ldots, \hat{k}_i, \ldots, k_{m+1}),$$

où le signe $\hat{}$ exprime que l'élément correspondant k_i doit être omis, définit une cochaîne $\delta c \in C^{m+1}(K; G)$. (En effet, si l'on permute deux sommets voisins k_a et k_{a+1}, $0 \leqslant a \leqslant m$, tous les termes de la somme (2) changent de signe à l'exception des termes d'indices a et $a+1$ qui sont permutés, car ils figurent dans cette somme avec des signes différents.)

Il est clair que l'application

$$\delta : C^m(K; G) \to C^{m+1}(K; G), \qquad c \mapsto \delta c,$$

est un homomorphisme qui est appelé *opérateur cobord.*

L'opérateur δ étant défini pour tout $m \geqslant 0$, on peut former la suite

$$C^\cdot(K; G) : C^0(K; G) \xrightarrow{\delta} C^1(K; G) \to \ldots$$
$$\ldots \to C^m(K; G) \xrightarrow{\delta} C^{m+1}(K; G) \to \ldots$$

Proposition 1. *La suite $C^\cdot(K; G)$ est un complexe de cochaînes, i.e.*

$$\delta(\delta c) = 0$$

pour toute cochaîne $c \in C^m(K; G)$ et tout $m \geqslant 0$.

Démonstration. Pour tout simplexe ordonné $(k_0, \ldots, k_{m+2}) \in K_{m+2}$ l'élément $\delta(\delta c)(k_0, \ldots, k_{m+2})$ du groupe G est la somme d'éléments de la forme

$$(3) \qquad c(k_0, \ldots, \hat{k}_i, \ldots, \hat{k}_j, \ldots, k_{m+2}), \quad 0 \leqslant i < j \leqslant m+2.$$

apparaissant deux fois chacun : une fois dans le calcul du terme $(-1)^j(\delta c)(k_0, \ldots, \hat{k}_j, \ldots, k_{m+2})$, l'autre fois dans celui du terme $(-1)^i(\delta c)(k_0, \ldots, \hat{k}_i, \ldots, k_{m+2})$. Ceci étant, dans le premier cas, le terme (3) sera précédé du signe $(-1)^{i+j}$ (puisque le sommet k_i est le i-ième sommet du simplexe $(k_0, \ldots, \hat{k}_j, \ldots, k_{m+2})$) et dans le second cas, du signe $(-1)^{i+j-1}$ (puisque le sommet k_j est $(j-1)$-ième sommet du simplexe $(k_0, \ldots, \hat{k}_i, \ldots, k_{m+2})$). Donc, tous les termes (3) se simplifient. □

Les cochaînes $c \in C^m(K; G)$ telles que $\delta c = 0$ s'appellent *cocycles* du schéma K sur le groupe G, et les cochaînes δc, où $c \in C^{m-1}(K; G)$, *cobords.* Les cocycles forment un sous-groupe $Z^m(K; G)$ (le

noyau de l'homomorphisme $\delta: C^m(K; G) \to C^{m+1}(K; G)$) et les cobords, un sous-groups $B^m(K; G)$ (l'image de l'homomorphisme $\delta: C^{m-1}(K; G) \to C^m(K; G)$; pour $m = 0$ on convient que $B^0(K; G) = 0$). On a ainsi défini le groupe quotient

$$H^m(K; G) = Z^m(K; G)/B^m(K; G)$$

puisque $B^m(K; G) \subset Z^m(K; G)$.

Ce groupe quotient s'appelle *m-ième groupe de cohomologie du schéma K sur le groupe G*, et ses éléments, *classes de cohomologie*.

Comparer avec les définitions générales du début de la leçon 20.

Remarque 3. On démontre (ce théorème est difficile!) que les groupes de cohomologie (et les groupes d'homologie dualement définis, cf. leçon 29 plus bas) de deux schémas sont isomorphes si les réalisations géométriques de ces schémas sont homéomorphes. Ceci permet d'appliquer l'appareil des groupes de cohomologie (et d'homologie) des schémas simpliciaux à l'étude des propriétés topologiques des polyèdres. La branche qui s'occupe de ces problèmes s'appelle t o p o l o g i e c o m b i n a t o i r e. Il y a un demi-siècle de cela, cette branche était autonome, mais aujourd'hui elle a été totalement absorbée par la t o p o l o g i e a l g é b r i q u e plus générale. [Le terme « topologie combinatoire » est parfois utilisé pour désigner toute la topologie algébrique, mais aujourd'hui cette situation ne correspond plus à la réalité.]

* * *

Lorsque K est le nerf d'un recouvrement $\mathfrak{U} = \{U_k, k \in K\}$ d'une variété différentiable $\mathscr{X}$, on peut généraliser la construction du complexe de cochaînes $C^\cdot(K; G)$ (qui est alors désigné par $C^\cdot(\mathfrak{U}; G)$). Supposons pour fixer les idées que $G = \mathbb{R}$.

On appelle *cochaîne de dimension $p \geqslant 0$ d'un recouvrement $\mathfrak{U}$ à valeurs dans un ensemble de fonctions* une application c définie sur K_p et associant à tout simplexe ordonné $(k_0, \ldots, k_p) \in K_p$ du nerf K de $\mathfrak{U}$ une fonction différentiable

$$c(k_0, \ldots, k_p) : U_{k_0} \cap \ldots \cap U_{k_p} \to \mathbb{R}$$

définie sur un ensemble ouvert (non vide!) $U_k \cap \ldots \cap U_{k_p}$ et satisfaisant à la condition d'antisymétrie (1). Ces cochaînes forment un espace vectoriel que l'on désignera par $C^{p,0}(\mathfrak{U})$. Le cobord $\delta c \in C^{p+1,0}(\mathfrak{U})$ d'une cochaîne $c \in C^{p,0}(\mathfrak{U})$ est défini par la formule (2) précédente (dans laquelle m est remplacé par p) à la seule différence que chaque terme $c(k_0, \ldots, \hat{k}_i, \ldots, k_{p+1})$ du second membre est supposé borné sur $U_{k_0} \cap \ldots \cap U_{k_{p+1}}$ (sinon la formule (2) n'aura pas de sens). Il est clair que la relation $\delta(\delta c) = 0$ (avec sa démonstration) reste en vigueur dans ce cas aussi, i.e. *la*

famille $C^{\cdot,0}(\mathfrak{U}) = \{C^{p,0}(\mathfrak{U}), \delta\}$ *des espaces vectoriels* $C^{p,0}(\mathfrak{U})$ *et des homomorphismes*

$$\delta : C^{p,0}(\mathfrak{U}) \to C^{p+1,0}(\mathfrak{U})$$

est un complexe de cochaînes.

De façon plus générale, pour tout $q \geqslant 0$ on peut introduire le complexe de cochaînes

$$C^{\cdot,q}(\mathfrak{U}) = \{C^{p,q}(\mathfrak{U}), \delta\},$$

composé de *cochaînes dont les valeurs sont des formes de degré* q, i.e. d'applications c définies sur K_p, satisfaisant à la condition d'antisymétrie (1) et dont les valeurs $c(k_0, \ldots, k_p)$ sont des formes de de degré q définies sur $U_{k_0} \cap \ldots \cap U_{k_p}$ (des éléments de l'espace vectoriel $\Omega^q(U_{k_0} \cap \ldots \cap U_{k_p})$). L'opérateur δ est défini dans ce cas par la même formule (2) (avec les mêmes précisions concernant les opérateurs de restriction). Pour $q = 0$ nous obtenons le complexe précédent $C^{\cdot,0}(\mathfrak{U})$.

A toute cochaîne $c \in C^{p,q}(\mathfrak{U})$, on peut associer une cochaîne $dc \in C^{p,q+1}(\mathfrak{U})$ en convenant par définition que

$$(\mathrm{d}c)(k_0, \ldots, k_p) = dc(k_0, \ldots, k_p)$$

pour tout simplexe ordonné $(k_0, \ldots, k_p) \in K_p$.

Il est clair que les applications

$$\mathrm{d} : C^{p,q}(\mathfrak{U}) \to C^{p,q+1}(\mathfrak{U}) \tag{4}$$

commutent aux opérateurs δ, i.e. pour tous $p, q \geqslant 0$ le diagramme

$$\begin{array}{ccc} C^{p,q}(\mathfrak{U}) & \xrightarrow{\delta} & C^{p+1,q}(\mathfrak{U}) \\ \mathrm{d}\downarrow & & \downarrow\mathrm{d} \\ C^{p,q+1}(\mathfrak{U}) & \xrightarrow{\delta} & C^{p+1,q+1}(\mathfrak{U}) \end{array}$$

est commutatif. Par définition (cf. définition 2 de la leçon 20) ceci exprime que les applications (4) forment une application de cochaînes

$$\mathrm{d} : C^{\cdot,q}(\mathfrak{U}) \to C^{\cdot,q+1}(\mathfrak{U})$$

du complexe $C^{\cdot,q}(\mathfrak{U})$ dans le complexe $C^{\cdot,q+1}(\mathfrak{U})$.

On remarquera que $\mathrm{d} \circ \mathrm{d} = 0$, i.e. pour tout $p \geqslant 0$ la famille $C^{p,\cdot}(\mathfrak{U}) = \{C^{p,q}(\mathfrak{U}), \mathrm{d}\}$ est un complexe de cochaînes. De plus, les applications

$$\delta : C^{p,q}(\mathfrak{U}) \to C^{p+1,q}(\mathfrak{U})$$

formeront une application de cochaînes du complexe $C^{p,\cdot}(\mathfrak{U})$ dans le complexe $C^{p+1,\cdot}(\mathfrak{U})$.

Définition 3. Une famille $C^{\cdot,\cdot} = \{C^{p,q}, \delta, \mathrm{d}\}$ de groupes $C^{p,q}$, où $p, q \geqslant 0$, et d'applications

$$\delta : C^{p,q} \to C^{p+1,q}, \quad \mathrm{d} : C^{p,q} \to C^{p,q+1}$$

s'appelle *complexe double* si

$$\delta \circ \delta = 0, \quad \mathrm{d} \circ \mathrm{d} = 0 \quad \text{et} \quad \mathrm{d} \circ \delta = \delta \circ \mathrm{d},$$

i.e. si

1° pour tout $q \geqslant 0$ (resp. $p \geqslant 0$) la famille $C^{\cdot,q} = \{C^{p,q}, \delta\}$ (resp. $C^{p,\cdot} = \{C^{p,q}, \mathrm{d}\}$) est un complexe de cochaînes;

2° les applications d forment une application de cochaînes

$$\mathrm{d} : C^{\cdot,q} \to C^{\cdot,q+1}$$

du complexe $C^{\cdot,q}$ dans le complexe $C^{\cdot,q+1}$, et par suite, les applications δ forment une application de cochaînes

$$\delta : C^{p,\cdot} \to C^{p+1,\cdot}$$

du complexe $C^{p,\cdot}$ dans le complexe $C^{p+1,\cdot}$.

On voit donc que pour tout recouvrement $\mathfrak{U}$ de la variété différentiable $\mathscr{X}$ on a construit le complexe double

$$C^{\cdot,\cdot}(\mathfrak{U}) = \{C^{p,q}(\mathfrak{U}), \delta, \mathrm{d}\}$$

que l'on appellera *complexe double du recouvrement* $\mathfrak{U}$.

* * *

Les complexes doubles seront représentés par des tables de la forme

$$\begin{tikzcd}
\vdots & \vdots & & \vdots & \vdots & \\
C^{0,q+1} \arrow[u] \arrow[r] & C^{1,q+1} \arrow[u] \arrow[r] & \dots \arrow[r] & C^{p,q+1} \arrow[u] \arrow[r, "\delta"] & C^{p+1,q+1} \arrow[u] \arrow[r] & \dots \\
C^{0,q} \arrow[u] \arrow[r] & C^{1,q} \arrow[u] \arrow[r] & \dots \arrow[r] & C^{p,q} \arrow[u, "\mathrm{d}"] \arrow[r, "\delta"] & C^{p+1,q} \arrow[u, "\mathrm{d}"] \arrow[r] & \dots \\
\vdots \arrow[u] & \vdots \arrow[u] & & \vdots \arrow[u] & \vdots \arrow[u] & \\
C^{0,0} \arrow[u] \arrow[r] & C^{1,0} \arrow[u] \arrow[r] & \dots \arrow[r] & C^{p,0} \arrow[u] \arrow[r] & C^{p+1,0} \arrow[u] \arrow[r] & \dots
\end{tikzcd}$$

Les complexes $C^{\cdot,q}$ seront en conséquence appelés *lignes* du complexe double $C^{\cdot,\cdot}$, et les complexes $C^{p,\cdot}$, ses *colonnes*.

En posant

$$C^m = C^{0,m} \oplus C^{1,m-1} \oplus \dots \oplus C^{m,0}$$

pour tout $m \geqslant 0$, on définit l'application

$$\mathfrak{d} : C^m \to C^{m+1}$$

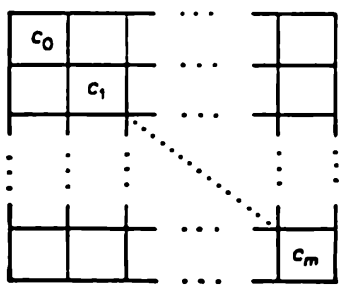
par la formule

$$\mathfrak{d}c = \delta c + \mathrm{d}'c,$$

où

$$\mathrm{d}'c = -(-1)^p \, \mathrm{d}c \quad \text{si} \quad c \in C^{p,\, m-p}.$$

Donc, les éléments du groupe G^m sont des chaînettes $c = (c_0, \ldots, c_m)$ des cochaînes du complexe $C^{\cdot\cdot}$, situées sur sa m-ième antidiagonale



et l'application $\mathfrak{d}$ associe à une telle chaînette une chaînette $\delta c = (e_0, \ldots, e_{m+1})$ telle que

$$(6) \qquad \begin{aligned} e_0 &= \mathrm{d}'c_0, \\ e_1 &= \delta c_0 + \mathrm{d}'c_1, \\ &\cdots\cdots\cdots\cdots \\ e_m &= \delta c_{m-1} + \mathrm{d}'c_m, \\ e_{m+1} &= \delta c_m, \end{aligned} \qquad \begin{tikzcd} & e_0 & & \\ c_0 \arrow[r] & e_1 & & \\ & c_1 \arrow[u] \arrow[r] & \cdots & \\ & & \vdots \arrow[u] & \\ & & c_{m-1} \arrow[r] & e_m \\ & & & c_m \arrow[u] \arrow[r] & e_{m+1}. \end{tikzcd}$$

En particulier, $\mathfrak{d}c = 0$ si et seulement si

$$(7) \qquad \begin{aligned} \mathrm{d}'c_0 &= 0, \\ \delta c_0 + \mathrm{d}'c_1 &= 0, \\ &\cdots\cdots\cdots \\ \delta c_{m-1} + \mathrm{d}'c_m &= 0, \\ \delta c_m &= 0, \end{aligned} \qquad \begin{tikzcd} & 0 & & \\ c_0 \arrow[r] & 0 & & \\ & c_1 \arrow[u] \arrow[r] & \cdots & \\ & & \vdots \arrow[u] & \\ & & c_{m-1} \arrow[r] & 0 \\ & & & c_m \arrow[u] \arrow[r] & 0. \end{tikzcd}$$

Puisque $d'\delta c = (-1)^{p+1} d\,\delta c$ et $\delta d' c = (-1)^p\,\delta\, dc$, il vient $d' \circ \delta + \delta \circ d' = 0$ et par suite,

$$\mathfrak{d} \circ \mathfrak{d} = (\delta + d') \circ (\delta + d') = \delta \circ \delta + \delta \circ d' + d' \circ \delta + d' \circ d' = 0.$$

Ceci exprime que *la famille* $\{C^m, \mathfrak{d}\}$ *est un complexe de cochaînes.*

Les groupes de cohomologie de ce complexe de cochaînes seront appelés *groupes de cohomologie du complexe double* $C^{\cdot\cdot\cdot}$ et désignés par $H^m(C^{\cdot\cdot\cdot})$.

Dans le cas particulier $C^{\cdot\cdot\cdot} = C^{\cdot\cdot\cdot}(\mathfrak{U})$, ces groupes seront appelés *groupes de cohomologie du recouvrement* $\mathfrak{U}$ et désignés par $H^m(\mathfrak{U})$.

* * *

Le complexe double du recouvrement $\mathfrak{U} = \{U_k, k \in K\}$ sera bordé de façon naturelle à gauche par le complexe de de Rham $\{\Omega^q \mathcal{X}, d\}$ de la variété $\mathcal{X}$ et, inférieurement, par le complexe $\{C^p(\mathfrak{U}); \mathbb{R}), \delta\}$ de cochaînes du recouvrement $\mathfrak{U}$ à coefficients dans le corps $\mathbb{R}$:

$$(8)\qquad \begin{tikzcd}
 & \cdots & & & & & \\
 & \vdots & \vdots & & & \vdots & \vdots \\
0 \arrow[r] & \Omega^q \mathcal{X} \arrow[u] \arrow[r, "j"] & C^{0,q}(\mathfrak{U}) \arrow[u] \arrow[r] & \cdots & \cdots \arrow[r] & C^{p,q}(\mathfrak{U}) \arrow[u, "d"] \arrow[r, "\delta"'] & C^{p+1,q}(\mathfrak{U}) \arrow[u, "d"] \arrow[r] & \cdots \\
 & \vdots \arrow[u] & \vdots \arrow[u] & \vdots & & \vdots \arrow[u] & \vdots \arrow[u] \\
0 \arrow[r] & \Omega^1 \mathcal{X} \arrow[u] \arrow[r, "j"] & C^{0,1}(\mathfrak{U}) \arrow[u] \arrow[r] & C^{1,1}(\mathfrak{U}) \arrow[u] \arrow[r] & \cdots & \vdots \arrow[u] & \vdots \arrow[u] \\
0 \arrow[r] & \Omega^0 \mathcal{X} \arrow[u] \arrow[r, "j"] & C^{0,0}(\mathfrak{U}) \arrow[u] \arrow[r] & C^{1,0}(\mathfrak{U}) \arrow[u] \arrow[r] & \cdots \arrow[r] & C^{p,0}(\mathfrak{U}) \arrow[u] \arrow[r] & C^{p+1,0}(\mathfrak{U}) \arrow[u] \arrow[r] & \cdots \\
 & & C^0(\mathfrak{U}; \mathbb{R}) \arrow[u, "i"] \arrow[r] & C^1(\mathfrak{U}; \mathbb{R}) \arrow[u, "i"] \arrow[r] & \cdots \arrow[r] & C^p(\mathfrak{U}; \mathbb{R}) \arrow[u, "i"] \arrow[r] & C^{p+1}(\mathfrak{U}; \mathbb{R}) \arrow[u, "i"] \arrow[r] & \cdots \\
 & & 0 \arrow[u] & 0 \arrow[u] & & 0 \arrow[u] & 0 \arrow[u]
\end{tikzcd}$$

où i sont les injections engendrées par l'identification d'un nombre $\lambda \in \mathbb{R}$ avec une fonction constante prenant la valeur λ, et les applications j sont définies par la forumle

$$(j\omega)(k) = \omega\,|_{U_k},$$

où $\omega \in \Omega^q \mathcal{X}$ et $k \in K$.

Il est clair que les applications i et j commutent respectivement aux opérateurs δ et d, i.e. forment les applications de cochaînes

$$i : C^{\cdot}(\mathfrak{U}; \mathbb{R}) \to C^{\cdot,0}(\mathfrak{U}), \quad j : \Omega^{\cdot}\mathcal{X} \to C^{0,\cdot}(\mathfrak{U}).$$

Ce sont en outre des monomorphismes qui vérifient les relations

$$d \circ i = 0, \quad \delta \circ j = 0.$$

(La relation $d \circ i = 0$ résulte du fait que si $f = \text{const}$, alors $df = 0$, et la relation $\delta \circ j = 0$, du fait que par définition

$$(\delta c)(k_0, k_1) = c(k_1) - c(k_0) \quad \text{sur } U_{k_0} \cap U_{k_1}$$

pour tout simplexe ordonné $(k_0, k_1) \in K_1$, et donc, si $c(k) = \omega$ sur U_k, alors $(\delta c)(k_0, k_1) = 0$.)

Cette situation mérite une définition spéciale.

Définition 4. On appelle *complexe double bordé* la table composée de groupes et d'homomorphismes

(9)
$$\begin{tikzcd}
 & \vdots & \vdots & & \vdots & \vdots \\
0 \arrow[r] & B^q \arrow[r, "j"] \arrow[u, "d"] & C^{0,q} \arrow[r] \arrow[u] & \cdots \arrow[r] & C^{p,q} \arrow[r, "\delta"] \arrow[u, "d"] & C^{p+1,q} \arrow[r] \arrow[u] & \cdots \\
0 \arrow[r] & B^1 \arrow[r, "j"] \arrow[u] & C^{0,1} \arrow[r] \arrow[u] & C^{1,1} \arrow[r] \arrow[u] & \cdots \arrow[r] & C^{p,1} \arrow[r] \arrow[u] & C^{p+1,1} \arrow[r] \arrow[u] & \cdots \\
0 \arrow[r] & B^0 \arrow[r, "j"] \arrow[u] & C^{0,0} \arrow[r] \arrow[u] & C^{1,0} \arrow[r] \arrow[u] & \cdots \arrow[r] & C^{p,0} \arrow[r] \arrow[u] & C^{p+1,0} \arrow[r] \arrow[u] & \cdots \\
 & & A^0 \arrow[r] \arrow[u, "i"] & A^1 \arrow[r] \arrow[u, "i"] & \cdots \arrow[r] & A^p \arrow[r, "\delta"] \arrow[u, "i"] & A^{p+1} \arrow[r] \arrow[u, "i"] & \cdots \\
 & & 0 \arrow[u] & 0 \arrow[u] & & 0 \arrow[u] & 0 \arrow[u] &
\end{tikzcd}$$

où la partie encadrée est un complexe double au sens de la définition 3 et où

1° les familles $A^\cdot = \{A^p, \delta\}$ et $B^\cdot = \{B^q, d\}$ sont des complexes de cochaînes ;

2° les applications i forment une application de cochaînes du complexe $A^\cdot$ dans le complexe $C^{\cdot,0}$, telle que

(10) $$d \circ i = 0;$$

3° les applications j forment une application de cochaînes du complexe $B^\cdot$ dans le complexe $C^{0,\cdot}$, telle que

$$\delta \circ j = 0;$$

4° toutes les applications i et j sont des monomorphismes.

* * *

De la relation (10) il résulte immédiatement que pour tout élément $a \in A$ l'identification de l'élément $ia \in C^{m,0}$ à la chaîne $(0, \ldots, 0, ia) \in C^m$ nous donne une application de cochaînes i de $A^\cdot$ dans $\{C^m, \delta\}$. [En effet, vu que $(d \circ i)\, a = 0$, il vient $\delta\,(ia) = (0, \ldots, 0, (\delta \circ i)\, a) = (0, \ldots, 0, i\,(\delta a)) = i\,(\delta a)$.] Les homomorphismes

(11) $$i^* : H^m(A^\cdot) \to H^m\ (C^{\cdot,\cdot})$$

induits par cette application de cochaînes seront appelés *homomorphismes bords inférieurs.*

On définit de façon analogue les *homomorphismes bords gauches*

$$j^* : H^m(B^\cdot) \to H^m(C^{\cdot\cdot}). \tag{12}$$

Pour un complexe double $C^{\cdot,\cdot}(\mathfrak{U})$ d'un recouvrement $\mathfrak{U}$ de $\mathcal{X}$ les homomorphismes bords inférieurs sont de la forme

$$i^* : H^m(\mathfrak{U};\ \mathbb{R}) \to H^m(\mathfrak{U}),$$

et les homomorphismes bords gauches, de la forme

$$j^* : H^m\mathcal{X} \to H^m(\mathfrak{U}).$$

* * *

On dira qu'un complexe double bordé (9) est *acyclique par rapport aux colonnes* si chacune de ses colonnes est une suite exacte, i.e. pour tous $p,\ q \geqslant 0$ on a

$$\operatorname{Ker}(\mathrm{d} : C^{p,q} \to C^{p,q+1}) = \begin{cases} \operatorname{Im}(i : A^p \to C^{p,0}) & \text{si } q = 0, \\ \operatorname{Im}(\mathrm{d} : C^{p,q-1} \to C^{p,q}) & \text{si } q > 0, \end{cases}$$

[Pour $q = 0$ cette égalité exprime que le monomorphisme i applique isomorphiquement le groupe A^p sur le groupe $H^0(C^{p,\cdot}) = \operatorname{Ker}(\mathrm{d}: C^{p,0} \to C^{p,1})$ du complexe $C^{p,\cdot}$, et pour $q > 0$, que le groupe $H^q(C^{p,\cdot})$ de cohomologie de ce complexe est égal à 0.]

On définit de façon analogue les complexes doubles *acycliques par rapport aux lignes*, pour lesquels

$$\operatorname{Ker}(\delta : C^{p,q} \to C^{p+1,q}) = \begin{cases} \operatorname{Im}(j : B^q \to C^{0,q}) & \text{si } p = 0, \\ \operatorname{Im}(\delta : C^{p-1,q} \to C^{p,q}) & \text{si } p > 0. \end{cases}$$

La clé du calcul des groupes de cohomologie de de Rham des variétés différentiables est fournie par le lemme algébrique suivant:

Lemme 1. *Si un complexe double bordé* (9) *est acyclique par rapport aux colonnes* (*resp. aux lignes*), *pour tout* $m \geqslant 0$ *l'homomorphisme bord inférieur* (11) (*resp. l'homomorphisme bord gauche* (12)) *est un isomorphisme.*

D é m o n s t r a t i o n. Supposons qu'un cocycle $a \in A^m$ du complexe $A^\cdot$ est tel que $i^*[a] = 0$, i.e. il existe une chaînette $c = (c_0, \ldots, c_{m-1}) \in C^{m-1}$ telle que $\mathrm{b}c = ia$. Par définition, ceci

exprime que

$$
(13)\qquad
\begin{aligned}
d'c_0 &= 0,\\
\delta c_0 + d'c_1 &= 0,\\
&\cdots\cdots\cdots\\
\delta c_{m-2} + d'c_{m-1} &= 0,\\
\delta c_{m-1} &= ia,
\end{aligned}
\qquad
\begin{tikzcd}
0 & & & & \\
c_0 \arrow[u] \arrow[r] & 0 & & & \\
& c_1 \arrow[u] \arrow[r] & \vdots & & \\
& & \cdots & & \\
& & c_{m-2} \arrow[u] \arrow[r] & 0 & \\
& & & c_{m-1} \arrow[u] \arrow[r] & ia.
\end{tikzcd}
$$

La première colonne du complexe $C^{\cdot\cdot\cdot}$ étant par hypothèse une suite exacte, il existe une cochaîne $b_0 \in C^{0,m-1}$ telle que $db_0 = c_0$. La cochaîne $c' = c - \delta b_0$ (i.e. plus exactement la cochaîne $c - \delta(b_0, 0, \ldots, 0)$) sera de la forme $(0, c'_1, \ldots, c'_{m-1})$ et vérifiera comme précédemment la relation $\delta c' = ia$. Donc, sans perdre en généralité, on peut d'emblée admettre que $c_0 = 0$.

La deuxième relation (13) sera alors de la forme $dc_1 = 0$ et, puisque la deuxième colonne du complexe $C^{\cdot\cdot\cdot}$ est une suite exacte, il existera un élément $b_1 \in C^{1,m-1}$ tel que $db_1 = c_1$. En remplaçant donc c par $c - \delta b_1$, on fait en sorte que $c_0 = 0$ et $c_1 = 0$ (tout en conservant la relation $\delta c = ia$).

En se déplaçant ainsi le long de l'antidiagonale, on arrive finalement à une chaînette c de la forme $(0, \ldots, 0, c_{m-1})$, où c_{m-1} est une cochaîne de $C^{m-1,0}$, telle que $dc_{m-1} = 0$ et $\delta c_{m-1} = ia$. Mais puisque $dc_{m-1} = 0$, il existe toujours en raison de l'exactitude de la colonne correspondante un élément $a' \in A^{m-1}$ tel que $ia' = c_{m-1}$ et donc tel que

$$i(\delta a') = \delta(ia') = \delta c_{m-1} = ia.$$

L'application i étant par hypothèse un monomorphisme, ceci prouve que $\delta a' = a$, i.e. que $[a] = 0$ dans $H^m(A^{\cdot})$. L'homomorphisme (11) est donc un monomorphisme.

Soit maintenant $c = (c_0, \ldots, c_{m-1})$ un cocycle à m dimensions du complexe $\{C^m, \delta\}$. Les relations (7) sont alors remplies pour les cochaînes $c_0, \ldots, c_m$. En particulier, $dc_0 = 0$. Donc, il existe de nouveau une cochaîne b_0 telle que la composante c_0 du cocycle $c - \delta b_0$ est nulle, donc, en vertu de la deuxième relation (7), $dc_1 = 0$ pour ce cocycle. Par conséquent, il existe une cochaîne b_1 telle que la composante c_1 du cocycle $c - db_0 - db_1$ sera nulle aussi.

En poursuivant ce processus, on obtient en définitive un cocycle cohomologue au cocycle donné, dont toutes les composantes $c_0, \ldots, c_{m-1}$ sont nulles et la composante $c_m = ia$, où a est un cocycle

du complexe $A^{\cdot}$, i.e. on obtient un cocycle qui est l'image du cocycle a par l'application $i: A^m \to C^m$. Ceci exprime que l'homomorphisme (11) est un épimorphisme.

On démontre *mutatis mutandis* la proposition analogue pour les complexes doubles acycliques par rapport aux lignes. □

Corollaire 1. *Si un complexe double bordé* (9) *est acyclique par rapport aux colonnes et aux lignes, pour tout* $m \geqslant 0$ *le groupe* $H^m(A^{\cdot})$ *est isomorphe au groupe* $H^m(B^{\cdot})$. □

Corollaire 2. *Si pour un recouvrement ouvert* $\mathfrak{U} = \{U_k, k \in K\}$ *d'une variété différentiable* $\mathscr{X}$ *le complexe bordé* (8) *est acyclique par rapport aux lignes et aux colonnes, pour tout* $m \geqslant 0$ *le groupe de cohomologie de de Rham* $H^m \mathscr{X}$ *de la variété* $\mathscr{X}$ *est isomorphe au groupe de cohomologie* $H^m(\mathfrak{U}; \mathbb{R})$ *du recouvrement* $\mathfrak{U}$:

$$H^m \mathscr{X} \approx H^m(\mathfrak{U}; \mathbb{R}). \ \square$$

Dans ce cas donc le groupe $H^m \mathscr{X}$ peut être calculé par un procédé purement combinatoire d'après le nerf K du *recouvrement* $\mathfrak{U}$. (Cf. calcul des groupes $H^m \mathbf{S}^n$ à la leçon précédente.)

* * *

La condition d'acyclicité du complexe double (8) par rapport aux lignes pour $p = 0$ exprime que chaque cocycle $c \in C^{0,q}(\mathfrak{U})$ est de la forme $i\omega$, où $\omega \in \Omega^q \mathscr{X}$, i.e. tel que $c(k) = \omega$ sur U_k pour tout $k \in K$. Or, il en va bien ainsi car la condition que la cochaîne $c: k \mapsto c(k) \in \Omega^q U_k$ est un cocycle exprime que

$$c(k_0) = c(k_1) \quad \text{sur} \quad U_{k_0} \cap U_{k_1}$$

pour tous indices $k_0, k_1 \in K$ (tels que $U_{k_0} \cap U_{k_1} \neq \varnothing$), i.e. les formes $c(k)$, $k \in K$, sont compatibles sur les intersections et constituent donc une forme $\omega \in \Omega^q \mathscr{X}$. Donc, *la condition d'acyclicité par rapport aux lignes pour* $p = 0$ *est remplie pour le complexe* (8) *d'un recouvrement arbitraire* $\mathfrak{U}$.

Les conditions assurant l'acyclicité du complexe (8) par rapport aux lignes (pour $p > 0$) et colonnes (pour $q \geqslant 0$) seront étudiées à la prochaine leçon.

LEÇON 22

Acyclicité par rapport aux lignes d'un complexe double d'un recouvrement numérotable. — Acyclicité par rapport aux colonnes d'un complexe double d'un recouvrement de Leray. — Théorème de de Rham-Leray. — Généralisation. — Groupes $E_2^{p,q}$. — Groupes $F^{p,q}$. — Groupe adjoint à un groupe gradué à filtration.

Définition 1. On dit qu'une famille $\{\eta_k\}$ de fonctions différentiables positives $\eta_k : \mathscr{X} \to \mathbb{R}$ est *localement finie* si tout point d'une variété $\mathscr{X}$ admet un voisinage en lequel seul un nombre fini de fonctions η_k sont non nulles. Une famille localement finie $\{\eta_k\}$ s'appelle *partition de l'unité* si

$$\sum_k \eta_k = 1$$

(cette somme a un sens en vertu de la condition de finitude locale). Une partition de l'unité $\{\eta_k\}$ est dite *subordonnée* à un recouvrement ouvert $\{U_k\}$ (avec le même ensemble d'indices) si $\eta_k = 0$ à l'extérieur de U_k pour tout k. (Signalons qu'il existe une autre définition plus restrictive qui implique que U_k contienne non seulement l'ensemble où $\eta_k \neq 0$ mais aussi son adhérence.) On dit qu'un recouvrement ouvert $\{U_k\}$ est *numérotable* s'il existe une partition de l'unité $\{\eta_k\}$ qui lui est subordonnée. Une variété $\mathscr{X}$ est *paracompacte* si chacun de ses recouvrements ouverts est numérotable.

Il se trouve que *si un recouvrement* $\mathfrak{U} = \{U_k;\ k \in K\}$ *d'une variété différentiable* $\mathscr{X}$ *est numérotable, son complexe double* $C^{\cdot\cdot}(\mathfrak{U})$ *est acyclique par rapport aux lignes.* En effet, en vertu de la remarque faite à la fin de la leçon précédente, il suffit seulement de prouver que pour tout $q \geqslant 0$, chaque cocycle c à p dimensions ($p > 0$) du complexe de cochaînes $C^{\cdot,q} = \{C^{p,q}(\mathfrak{U}), \delta\}$ est un cobord. A cet effet, associons à une cochaîne $c \in C^{p,q}(\mathfrak{U})$ la cochaîne $Dc \in C^{p-1,q}(\mathfrak{U})$ définie (visiblement de façon unique) par la formule

$$(1) \qquad (Dc)(k_0, \ldots, k_{p-1}) = \sum_{k \in K} \eta_k c(k, k_0, \ldots, k_{p-1}),$$

$$(k_0, \ldots, k_{p-1}) \in K_{p-1},$$

où $\{\eta_k\}$ est une partition de l'unité subordonnée au recouvrement $\mathfrak{U}$. Alors pour tout simplexe $(k_0, \ldots, k_p) \in K_p$

$$(\delta Dc)(k_0, \ldots, k_p) = \sum_{i=0}^{p} (-1)^i \sum_{k \in K} \eta_k c(k, k_0, \ldots, \hat{k}_i, \ldots, k_p) =$$

$$= \sum_{k \in K} \eta_k \sum_{i=0}^{p} (-1)^i c(k, k_0, \ldots, \hat{k}_i, \ldots, k_p)$$

et

$$(D\delta c)(k_0, \ldots, k_p) = \sum_{k \in K} \eta_k (\delta c)(k, k_0, \ldots, k_p) =$$

$$= \sum_{k \in K} \eta_k \left[c(k_0, \ldots, k_p) + \sum_{i=0}^{p} (-1)^{i+1} c(k, k_0, \ldots, \hat{k}_i, \ldots, k_p) \right] =$$

$$= c(k_0, \ldots, k_p) - \sum_{k \in K} \eta_k \sum_{i=0}^{p} (-1)^i c(k, k_0, \ldots, \hat{k}_i, \ldots, k_p) =$$

$$= c(k_0, \ldots, k_p) - (\delta Dc)(k_0, \ldots, k_p),$$

i.e.

$$D\delta c + \delta Dc = c$$

pour toute cochaîne c. Donc, si $\delta c = 0$, alors $c = \delta Dc$. □

* * *

Considérons maintenant les conditions assurant l'acyclicité du complexe $C^{\cdot\cdot}(\mathfrak{U})$ par rapport aux colonnes.

On rappelle (cf. définition 1 de la leçon 10) que $\prod_\alpha \mathscr{X}_\alpha$ désigne l'ensemble de toutes les familles $\{x_\alpha\}$, où $x_\alpha \in \mathscr{X}_\alpha$. Lorsque les $\mathscr{X}_\alpha$ sont des espaces vectoriels, cet ensemble est un espace vectoriel par rapport aux opérations $\{x_\alpha\} + \{y_\alpha\} = \{x_\alpha + y_\alpha\}$, $\lambda\{x_\alpha\} = \{\lambda x_\alpha\}$ et s'appelle *produit direct* des espaces vectoriels $\mathscr{X}_\alpha$. On définit de façon analogue les produits directs de groupes, de complexes de cochaînes, de complexes doubles, etc.

Ceci étant, il est clair que par exemple *les groupes de cohomologie d'un produit direct de complexes seront produits directs des groupes de cohomologie des facteurs.*

En comparant la définition générale du produit direct d'espaces vectoriels avec celle des espaces vectoriels $C^{p,q}(\mathfrak{U})$, on s'aperçoit immédiatement que *chaque espace vectoriel* $C^{p,q}(\mathfrak{U})$ *est un produit direct*

$$\prod_{\{k_0, \ldots, k_p\} \subset K} \Omega^q(U_{k_0} \cap \ldots \cap U_{k_p})$$

d'espaces vectoriels $\Omega^q(U_{k_0} \cap \ldots \cap U_{k_p})$ *par rapport aux simplexes à p dimensions* $\{k_0, \ldots, k_p\}$ *du schéma simplicial* K. De plus, l'application $d : C^{p,q}(\mathfrak{U}) \to C^{p,q+1}(\mathfrak{U})$ associe à tout facteur $\Omega^q(U_{k_0} \cap \ldots \cap U_{k_p})$ de l'espace vectoriel $C^{p,q}(\mathfrak{U})$ le facteur $\Omega^{q+1}(U_{k_0} \cap \ldots \cap U_{k_p})$ de l'espace vectoriel $C^{p,q+1}(\mathfrak{U})$ et n'est autre que la dérivation extérieure

$$d : \Omega^q(U_{k_0} \cap \ldots \cap U_{k_p}) \to \Omega^{q+1}(U_{k_0} \cap \ldots \cap U_{k_p}).$$

Ceci exprime que pour tout $p \geqslant 0$ le complexe $C^{p,\cdot}(\mathfrak{U})$ est un produit direct de complexes de de Rham $\Omega^{\cdot}(U_{k_0} \cap \ldots \cap U_{k_p})$, et donc ses groupes de cohomologie sont les produits directs des groupes $H^q(U_{k_0} \cap \ldots \cap U_{k_p})$. S'agissant donc du complexe $C^{\cdot\cdot}(\mathfrak{U})$, la condition d'acyclicité par rapport aux colonnes est remplie pour $q > 0$ si et seulement si les groupes de cohomologie $H^q(U_{k_0} \cap \ldots \cap U_{k_p})$, $q > 0$, sont nuls pour tout simplexe $\{k_0, \ldots, k_p) \subset K$. De façon analogue, pour tout $p \geqslant 0$ le groupe $C^p(\mathfrak{U}; \mathbb{R})$ n'étant autre que le produit direct d'espaces vectoriels $\mathbb{R}$ par rapport à tous les simplexes $\{k_0, \ldots, k_p) \subset K$, dire que le complexe $C^{\cdot\cdot}(\mathfrak{U})$ est acyclique par rapport aux colonnes pour $q = 0$ revient à dire que le groupe $H^0(U_{k_0} \cap \ldots \cap U_{k_p})$ est isomorphe à $\mathbb{R}$ pour tout simplexe $\{k_0, \ldots, k_p\} \subset K$.

Ces conditions étant visiblement remplies en vertu du lemme de Poincaré (et de la proposition 2 de la leçon 20) lorsque toutes les intersections $U_{k_0} \cap \ldots \cap U_{k_p}$ sont difféomorphes à une boule ouverte $\mathring{\mathbb{B}}^n$ de l'espace $\mathbb{R}^n$ (où comme toujours $n = \dim \mathscr{X}$), on voit que *le complexe* $C^{\cdot\cdot\cdot}(\mathfrak{U})$ *est acyclique par rapport à ses colonnes si pour tous indices* $k_0, \ldots, k_p \in K$ *l'intersection* $U_{k_0} \cap \ldots \cap U_{k_p}$, *pourvu qu'elle ne soit pas vide, est difféomorphe à la boule* $\mathring{\mathbb{B}}^n$.

Définition 2. On dira qu'un recouvrement ouvert $\mathfrak{U} = \{U_k; k \in K\}$ d'une variété différentiable $\mathscr{X}$ à n dimensions est un *recouvrement de Leray* si pour tous indices $k_0, \ldots, k_p \in K$ l'intersection $U_{k_0} \cap \ldots \cap U_{k_p}$ soit est vide, soit est difféomorphe à la boule $\mathring{\mathbb{B}}^n$.

Donc, *pour tout recouvrement de Leray* $\mathfrak{U}$, *le complexe* $C^{\cdot\cdot\cdot}(\mathfrak{U})$ *est acyclique par rapport à ses colonnes.*

* * *

En collationnant les résultats obtenus, on obtient grâce au corollaire 2 du lemme 1 de la leçon précédente le théorème suivant :

Théorème 1. *Pour tout recouvrement numérotable de Leray d'une variété différentiable* $\mathscr{X}$, *les groupes de cohomologie* $H^m(\mathfrak{U}; \mathbb{R})$ *sont*

isomorphes aux groupes de cohomologie de de Rham $H^m\mathscr{X}$ de la variété $\mathscr{X}$.

Ce théorème s'appelle théorème de de Rham-Leray.

On démontre sans peine la formule explicitant l'isomorphisme

$$H^m(\mathfrak{U};\ \mathbb{R}) \to H^m\mathscr{X}. \tag{2}$$

Exercice 1. Si l'on applique m fois l'opérateur $\mathrm{d} \circ D$, où D est un opérateur de $C^{p,q}(\mathfrak{U})$ dans $C^{p-1,q}(\mathfrak{U})$ défini par la formule (1), à la cochaîne $ic \in C^{m,0}(\mathfrak{U})$, $c \in C^m(\mathfrak{U};\ \mathbb{R})$, on obtient une cochaîne de $C^{0,m}(\mathfrak{U})$. Montrer que

1° lorsque c est un cocycle, la cochaîne $(\mathrm{d} \circ D)^m c$ est de la forme $j\omega$, où ω est une forme fermée de $\Omega^m\mathscr{X}$;

2° la classe de cohomologie $[\omega]$ de la forme ω dépend de la seule classe de cohomologie $[c]$ du cocycle c;

3° l'application $[c] \mapsto [\omega]$ est au signe près un isomorphisme (2).

Corollaire 1. *Pour tout recouvrement numérotable de Leray $\mathfrak{U}$ d'une variété différentiable $\mathscr{X}$, les groupes de cohomologie $H^m(\mathfrak{U};\ \mathbb{R})$ sont indépendants — à un isomorphisme près — de ce recouvrement et sont définis exclusivement par la variété $\mathscr{X}$.* □

Remarque 1. On démontre — par des raisonnements totalement différents ne dépendant pas des formes différentielles — que le corollaire 1 est encore valable pour les groupes de cohomologie $H^m(\mathfrak{U};\ G)$ sur un groupe quelconque G.

Définition 3. Les groupes $H^m(\mathfrak{U};\ G)$, où $\mathfrak{U}$ est un recouvrement numérotable de Leray de $\mathscr{X}$, s'appellent *groupes de cohomologie de Čech* de la variété $\mathscr{X}$ et sont désignés par $H^m(\mathscr{X};\ G)$. Ils sont définis si la variété $\mathscr{X}$ possède au moins un recouvrement numérotable de Leray.

Cette définition est cohérente en vertu de la remarque 1. Elle permet de reformuler le théorème 1 sous la forme suivante:

Théorème 1a. *Si une variété $\mathscr{X}$ possède un recouvrement numérotable de Leray, ses groupes de cohomologie de de Rham sont isomorphes à ses groupes de cohomologie de Čech:*

$$H^m\mathscr{X} \approx H^m(\mathscr{X};\ \mathbb{R}), \quad m \geqslant 0.$$

Bien que le problème de l'existence d'un recouvrement numérotable de Leray pour une variété donnée $\mathscr{X}$ ne pose en principe pas de difficultés, il est intéressant de l'envisager dans le cas général.

Théorème 2. *Toute variété séparée satisfaisant au deuxième axiome de dénombrabilité (en particulier, toute variété compacte séparée) est paracompacte, i.e. chacun de ses recouvrements ouverts est numérotable.*

Remarque 2. On démontre que, réciproquement, *toute variété séparée paracompacte connexe satisfait au deuxième axiome de dénombrabilité.*

Cf. plus bas les remarques 2 et 3 de la leçon 24.

Théorème 3. *Toute variété séparée paracompacte possède un recouvrement de Leray. Si une variété est compacte, elle possède un recouvrement fini de Leray.*

Corollaire 1. *Les groupes de cohomologie de Čech $H^m(\mathscr{X}; G)$ sont définis pour toute variété séparée paracompacte $\mathscr{X}$. Si la variété $\mathscr{X}$ est compacte, ces groupes possèdent un nombre fini de génératrices sur G (pour $G = \mathbb{R}$, ces groupes sont finidimensionnels).*

Corollaire 2. *Les groupes de cohomologie de de Rham $H^m\mathscr{X}$ d'une variété séparée paracompacte $\mathscr{X}$ sont isomorphes à ses groupes de Čech sur le corps $\mathbb{R}$:*

$$H^m\mathscr{X} \approx H^m(\mathscr{X}; \mathbb{R}).$$

Le corollaire 2 est généralement appelé théorème de de Rham pour les groupes de cohomologie de Čech.

[D'une façon générale, on appelle théorèmes de de Rham toute une série de théorèmes décrivant les groupes $H^m\mathscr{X}$ en termes topologiques. Ceci est dû au fait que le tout premier de ces théorèmes que nous verrons à la leçon 29 a été établi par de Rham.]

Corollaire 3. *Les espaces vectoriels $H^m\mathscr{X}$, où $\mathscr{X}$ est une variété séparée compacte, sont finidimensionnels (donc, les nombres de Betti $h^m\mathscr{X}$ sont définis).* □

On démontrera le théorème 2 à la leçon 24. S'agissant du théorème 3, nous sommes contraints, comme dans le cas du lemme de Gauss de la leçon 5, de l'ajourner au prochain ouvrage de cette série.

* * *

D'une façon générale, même pour les variétés $\mathscr{X}$ les plus simples, les recouvrements de Leray $\mathfrak{U}$ contiennent un nombre assez élevé d'éléments, et le calcul des groupes $H^m(\mathfrak{U}; \mathbb{R}) = H^m(\mathscr{X}; \mathbb{R})$, s'il n'est pas compliqué, devient en tout cas laborieux. D'où l'intérêt des méthodes de calcul des groupes $H^m(\mathscr{X}; \mathbb{R})$ opérant sur des recouvrements qui ne sont pas de Leray. (Les exemples de la leçon 20 montrent que cette approche peut être payante.) Ceci étant en vertu du théorème 2 (qu'à vrai dire nous n'avons pas encore démontré) on peut limiter notre étude des variétés, satisfaisant au deuxième axiome de dénombrabilité, aux seuls recouvrements numérotables $\mathfrak{U} = \{U_k, k \in K\}$ pour lesquels le complexe bordé $C^{\cdot\cdot\cdot}(\mathfrak{U})$ est, comme on le sait, acyclique par rapport à ses lignes, et donc, ses groupes de cohomologie $H^m(\mathfrak{U})$ sont isomorphes aux groupes de cohomologie de de Rham $H^m\mathscr{X}$ de la variété $\mathscr{X}$. Le calcul des groupes $H^m\mathscr{X}$ se ramène ainsi à celui des groupes $H^m(\mathfrak{U})$.

On admettra que sont connus:

a) le nerf K du recouvrement $\mathfrak{U}$ (i.e. les indices $k_0, \ldots, k_p \in K$ pour lesquels l'intersection $U_{k_0} \cap \ldots \cap U_{k_p}$ n'est pas vide);

b) les groupes de cohomologie de de Rham $H^q(U_{k_0} \cap \ldots \cap U_{k_p})$ de la sous-variété ouverte $U_{k_0} \cap \ldots \cap U_{k_p}$ pour tout simplexe $\{k_0, \ldots, k_p\}$ du nerf K; et on cherchera une procédure algébrique qui nous permette à défaut de calculer entièrement les groupes $H^m(\mathfrak{U}) \approx H^m \mathscr{X}$ (ce qui restera du domaine du rêve) de recueillir à la rigueur la plus ample information sur ces groupes.

* * *

Soit $C^{\cdot\cdot\cdot} = \{C^{p,q}; \delta, \mathrm{d}\}$ un complexe double, provisoirement arbitraire. Les colonnes $C^{p,\cdot} = \{C^{p,q}; \mathrm{d}\}$ du complexe $C^{\cdot\cdot\cdot}$ étant d'ordinaires complexes de cochaînes, leurs groupes de cohomologie $H^q(C^{p,\cdot})$ sont définis. Ces groupes sont traditionnellement désignés par $E_1^{p,q}$. (On attire l'attention du lecteur sur l'ordre des indices supérieurs !)

Les opérateurs cobords horizontaux $\delta : C^{p,q} \to C^{p+1,q}$ engendrent par définition une application de cochaînes $C^{p,\cdot} \to C^{p+1,\cdot}$, donc ils induisent les homomorphismes

$$\delta^* : E_1^{p,q} \to E_1^{p+1,q}. \tag{3}$$

Puisque $\delta \circ \delta = 0$, il vient $\delta^* \circ \delta^* = 0$ et, par suite, pour tout $q \geqslant 0$ la famille $E_1^{\cdot,q} = \{E_1^{p,q}; \delta^*\}$ de groupes et d'homomorphismes est un complexe de cochaînes. (Dans la suite on écrira tout simplement δ à la place de δ^*.) Les groupes de cohomologie $H^p(E_1^{\cdot,q})$ des complexes $E_1^{\cdot,q}$ seront aussi par tradition désignés par $E_2^{p,q}$ et représentés par une table de la forme

(4)

$E_2^{0,q}$	$E_2^{1,q}$	$\ldots$	$E_2^{p,q}$	$\ldots$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
$E_2^{0,1}$	$E_2^{1,1}$	$\ldots$	$E_2^{p,1}$	$\ldots$
$E_2^{0,0}$	$E_2^{1,0}$	$\ldots$	$E_2^{p,0}$	$\ldots$

Par convention, $E_2^{p,q} = H_\delta^p H_\mathrm{d}^q(C^{\cdot\cdot\cdot})$ (et $E_2^{p,q} = H_\delta^p H_\mathrm{d}^q(\mathfrak{U})$ si $C^{\cdot\cdot\cdot} = C^{\cdot\cdot\cdot}(\mathfrak{U})$), où les indices inférieurs désignent l'opérateur cobord par rapport auquel sont calculés les groupes de cohomologie.

Signalons que dans le cas du complexe double $C^{\cdot\cdot\cdot}(\mathfrak{U})$ *l'information contenue dans les points* a) *et* b) *ci-dessus suffit en principe pour calculer les groupes* $E_2^{p,q}$. En effet, on sait que chaque colonne $C^{p,\cdot}(\mathfrak{U})$ du complexe $C^{\cdot\cdot\cdot}(\mathfrak{U})$ est un produit direct de complexes de de Rham de la forme $\Omega^{\cdot}(U_{k_0} \cap \ldots \cap U_{k_p})$, donc, ses groupes de cohomologie $E_1^{p,q} = H^q(C^{p,\cdot}(\mathfrak{U}))$ sont des produits directs des groupes de cohomologie de de Rham $H^q(U_{k_0} \cap \ldots \cap U_{k_p})$. Ceci exprime que le passage du groupe $C^{p,q}(\mathfrak{U})$ au groupe $E_1^{p,q}$ consiste, si on se limite aux cochaînes

$$c : (k_0, \ldots, k_p) \mapsto c(k_0, \ldots, k_p) \in \Omega^q(U_{k_0} \cap \ldots \cap U_{k_p})$$

dont les valeurs $c(k_0, \ldots, k_p)$ sont des formes fermées, à passer à leurs classes de cohomologie

$$[c(k_0, \ldots, k_p)] \in H^q(U_{k_0} \cap \ldots \cap U_{k_p}).$$

On peut donc admettre que les éléments du groupe $E_1^{p,q}$ sont des fonctions antisymétriques

$$\bar{c} : (k_0, \ldots, k_p) \mapsto \bar{c}(k_0, \ldots, k_p),$$

définies sur K_p, telles que $\bar{c}(k_0, \ldots, k_p) \in H^q(U_{k_0} \cap \ldots \cap U_{k_p})$ (*des cochaînes à valeurs dans un groupe de cohomologie de de Rham*). Ceci étant, l'opérateur (3) sera défini par la formule

$$(\delta^* \bar{c})(k_0, \ldots, k_{p+1}) = \sum_{i=0}^{p+1} (-1)^i \bar{c}(k_0, \ldots, \hat{k}_i, \ldots, k_{p+1}),$$

où au second membre on a omis d'écrire les homomorphismes de restriction

$$H^q(U_{k_0} \cap \ldots \cap \hat{U}_{k_i} \cap \ldots \cap U_{k_p}) \to H^q(U_{k_0} \cap \ldots \cap U_{k_p})$$

(cf. formule (2) de la leçon précédente). Ceci nous fournit un algorithme de calcul des groupes $E_2^{p,q}$, qui est assez opérationnel si ce calcul n'est pas trop volumineux. □

Donc, nous pouvons considérer que les groupes $E_2^{p,q}$ pour les complexes doubles de la forme $C^{\cdot\cdot\cdot}(\mathfrak{U})$ sont connus du moins théoriquement.

Exemple 1. Si $\mathfrak{U} = \{U, V\}$ est le recouvrement à deux éléments de la sphère S^n envisagé à la leçon 20, parmi les groupes $H^m(U)$, $H^m(V)$, et $H^m(U \cap V)$, seuls sont non nuls les groupes

$$H^0(U) \approx \mathbb{R}, \quad H^0(V) \approx \mathbb{R}, \quad H^0(U \cap V) \approx \mathbb{R}$$

et

$$H^{n-1}(U \cap V) \approx H^{n-1}S^{n-1} \approx \mathbb{R}.$$

Donc, le complexe $E_1^{\cdot,0}$ est isomorphe au complexe $C^{\cdot}(\mathfrak{U}; \mathbb{R})$, et par conséquent, *les groupes* $E_2^{p,0}$ *sont isomorphes aux groupes* $H^p(\mathfrak{U}; \mathbb{R})$.

(C'est un fait général valable pour toute variété $\mathscr{X}$ et pour tout recouvrement $\mathfrak{U}$ de $\mathscr{X}$ dont les intersections $U_{k_0} \cap \ldots \cap U_{k_p}$ sont *connexes*.)

D'autre part, il est aisé de voir que $H^0(\mathfrak{U}; \mathbb{R}) \approx \mathbb{R}$ (autre fait général valable pour tout schéma simplicial connexe dans un sens clair) et que $H^p(\mathfrak{U}; \mathbb{R}) = 0$ pour $p > 0$ (le nerf du recouvrement $\mathfrak{U}$, ou plus exactement sa réalisation géométrique, est le segment

(5)

donc, pour $p > 1$, même le groupe $C^p(\mathfrak{U}; \mathbb{R})$ est nul, et pour $p = 1$, tout élément du groupe $C^1(\mathfrak{U}; \mathbb{R}) = Z^1(\mathfrak{U}; \mathbb{R})$ est défini de façon unique par le nombre $\lambda = c(0, 1)$, donc, est égal à δe, où e est une cochaîne de dimension zéro définie par la formule $e(0) = 0$, $e(1) = \lambda$).

Donc,

$$E_2^{p,0} = \begin{cases} \mathbb{R} \text{ si } p = 0, \\ 0 \text{ si } p > 0. \end{cases}$$

Parmi les groupes $E_1^{p,q}$, $q > 0$, seul est non nul le groupe $E_1^{1,n-1} = C^1(\mathfrak{U}; \mathbb{R})$ isomorphe à $\mathbb{R}$. Donc, parmi les groupes $E_2^{p,q}$, $q > 0$, seul sera aussi non nul le groupe $E_2^{1,n-1} \approx \mathbb{R}$.

Par conséquent, pour le recouvrement $\{U, V\}$ de S^n la table des groupes $E_2^{p,q}$ est de la forme

$n-1$		$\mathbb{R}$	...		...
	. . .	. . .	. . .	. . .	. . .
0	$\mathbb{R}$		...		...
	0	1			

(les autres cases contiennent des zéros). □

* * *

Il reste à passer des groupes $E_2^{p,q}$ aux groupes $H^m(C^{\cdot\cdot})$. Pour cela on décrira d'abord les groupes $E_2^{p,q}$ à l'aide des groupes $C^{p,q}$. Ceci étant, pour éviter les innombrables restrictions, on admettra que les groupes $C^{p,q}$ sont définis aussi pour les $p, q < 0$ et sont nuls dans ce cas:

$$C^{p,q} = 0 \quad \text{si} \quad p < 0 \quad \text{ou} \quad q < 0.$$

Par définition, chaque élément x du groupe $E_2^{p,q}$ est une classe de δ-cohomologie d'un δ-cocycle $\bar{c} \in E_1^{p,q}$ qui à son tour est une classe de d-cohomologie d'un d-cocycle $c \in C^{p,q}$. (Dans le complexe du

recouvrement $\mathfrak{U}$ l'élément $\bar{c}$ est un cocycle à valeurs dans un groupe de cohomologie de de Rham et c s'obtient à partir de $\bar{c}$ en choisissant un représentant dans chaque classe de cohomologie $\bar{c}$ $(k_0, \ldots, k_p)$.) Donc, chaque élément $x \in E_2^{p,q}$ est défini par un élément $c \in C^{p,q}$. On écrira $x = [c]_2$ (et $\bar{c} = [c]_1$).

Pour que l'élément $[c]_2$ soit défini, il est nécessaire bien sûr que $dc = 0$. En outre, l'élément $\delta c \in C^{p+1,q}$ doit être de la forme dc_1, où $c_1 \in C^{p+1,q-1}$ (pour que l'élément $[c]_1 \in E_1^{p,q}$ soit un δ-cocycle). Du reste il nous sera plus commode de remplacer d par d′ et de mettre δc sous la forme $-d'c_1$ (ce qui en principe est sans importance). Réciproquement, si $dc = 0$, l'élément $[c]_1$ est défini, et si $\delta c = -d'c_1$, alors $\delta [c]_1 = 0$, donc est défini l'élément $[c]_2$. Par conséquent, *l'élément* $[c]_2$ *est défini si et seulement s'il existe une chaînette à deux éléments* (c, c_1) *de la forme*

$$(6) \qquad \begin{array}{ccc} 0 & & \\ {\scriptstyle d'}\uparrow & & \\ c & \xrightarrow{\delta} & \\ & \uparrow{\scriptstyle d_1} & \\ & c_1, & \end{array} \qquad \begin{aligned} d'c &= 0, \\ \delta c + d'c_1 &= 0. \end{aligned}$$

On démontre de façon analogue que $[c]_2 = [c']_2$ *si et seulement s'il existe des cochaînes* $a \in C^{p-1,q}$ *et* $b \in C^{p,q-1}$ *telles que*

$$d'a = 0 \quad et \quad c' - c = \delta a - d'b;$$

schématiquement

$$(7) \qquad \begin{array}{ccccc} 0 & & & & \\ \uparrow & & & & \\ a & \rightarrow & c' & - & c \\ & & & & \uparrow \\ & & & & b. \end{array}$$

On sait d'autre part (cf. formule (7) de la leçon précédente) que les chaînettes de la forme

$$(8) \qquad \begin{array}{ccccccccccc} 0 & & & & & & & & & & \\ \uparrow & & & & & & & & & & \\ c_0 & \rightarrow & & & & & & & & & \\ & & \uparrow & & & & & & & & \\ & & c_1 & \rightarrow & & & & & & & \\ & & & & \uparrow & & & & & & \\ & & & & c_2 & \rightarrow & & & & & \\ & & & & & & \ddots & & & & \\ & & & & & & & \uparrow & & & \\ & & & & & & & c_{m-1} & \rightarrow & & \\ & & & & & & & & & \uparrow & \\ & & & & & & & & & c_m & \rightarrow 0, \end{array} \qquad \begin{aligned} d'c_0 &= 0, \\ \delta c_0 + d'c_1 &= 0, \\ \delta c_1 + d'c_2 &= 0, \\ &\cdots\cdots \\ \delta c_{m-1} + d'c_m &= 0, \\ \delta c_m &= 0, \end{aligned}$$

à longueur maximale (finissant sur la ligne inférieure du complexe $C^{\cdot\cdot\cdot}$) sont des cocycles du complexe $C^{\cdot\cdot\cdot}$ (par rapport à l'opérateur δ). Donc, si le cocycle (8) est tel que $c_0 = 0, \ldots, c_{p-1} = 0$, l'élément $[c_p]_2$ du groupe $E_2^{p,\,m-p}$ sera défini pour sa composante c_p.

Donc, la correspondance $c \mapsto [c_p]_2$ définit une application

$$Z^{p,\,m-p} \to E_2^{p,\,m-p}, \tag{9}$$

où $Z^{p,\,m-p}$ est un sous-groupe du groupe des δ-cocycles $Z^m(C^{\cdot\cdot\cdot})$ du complexe double $C^{\cdot\cdot\cdot}$, sous-groupe qui est composé des cocycles $c = (c_0, \ldots, c_m)$ pour lesquels $c_0 = 0, \ldots, c_{p-1} = 0$. Ceci étant, la définition de l'opérateur δ (cf. formule (6) de la leçon précédente) entraîne immédiatement que, si les cocycles $c = (0, \ldots, 0, c_p, \ldots, c_m)$ et $c' = (0, \ldots, 0, c'_p, \ldots, c'_m)$ du groupe $Z^{p,\,m-p}$ sont cohomologues, il existe une chaînette $(a_0, a_1, \ldots, a_p, 0, \ldots, 0)$ telle que

$$\begin{aligned} d'a_0 &= 0, \\ \delta a_1 + d'a_1 &= 0, \\ &\ldots\ldots\ldots\ldots \\ \delta a_{p-2} + d'a_{p-1} &= 0, \\ \delta a_{p-1} + d'a_p &= c'_p - c_p, \end{aligned} \tag{10}$$

$$\begin{tikzcd} 0 & & & & & \\ a_0 \arrow[u] \arrow[r] & & & & & \\ & a_1 \arrow[u] \arrow[r] & & & & \\ & & \ddots & & & \\ & & & a_{p-2} \arrow[u] \arrow[r] & & \\ & & & & a_{p-1} \arrow[u] \arrow[r] & c'_p - c_p \\ & & & & & a_p \arrow[u] \end{tikzcd}$$

Cette relation se transforme en une relation exprimée par le diagramme (7) pour $a_0 = 0, \ldots, a_{p-2} = 0$, mais dans le cas général elle est plus faible que la relation (7). Ceci nous oblige à envisager le groupe quotient $\overline{E}_2^{p,\,m-p}$ du groupe $E_2^{p,\,m-p}$ par le sous-groupe de tous les éléments de la forme $[\delta a_{p-1} + d'a_p]_2$ et le composé

$$Z^{p,\,m-p} \to \overline{E}_2^{p,\,m-p} \tag{11}$$

de l'homomorphisme (9) et de l'homomorphisme de passage au quotient $E_2^{p,\,m-p} \to \overline{E}_2^{p,\,m-p}$. Contrairement à l'homomorphisme (9), l'homomorphisme (11) envoie tous les cobords contenus dans $Z^{p,\,m-p}$ en 0, et, par suite, induit un homomorphisme

$$F^{p,\,m-p} \to \overline{E}_2^{p,\,m-p}, \tag{12}$$

où $F^{p,\,m-p}$ est le sous-groupe de $H^m(C^{\cdot\cdot\cdot})$, image du sous-groupe $Z^{p,\,m-p}$ par l'homomorphisme de passage au quotient $Z^m(C^{\cdot\cdot\cdot}) \to$

$\to H^m(C^{\cdot\cdot\cdot})$, i.e. est composé des classes de cohomologie $[c]$ contenant le cocycle c de la forme $(0, \ldots, 0, c_p, \ldots, c_m)$.

Il est clair que le sous-groupe $F^{p+1,\, m-p-1}$ du groupe $F^{p,\, m-p}$ est envoyé dans 0 par l'homomorphisme (12). Donc, cet homomorphisme induit un homomorphisme

$$F^{p,\, m-p}/F^{p+1,\, m-p-1} \to \overline{E}_2^{p,\, m-p}. \tag{13}$$

D'autre part, dire que la classe de cohomologie $[c] \in F^{p,\, m-p}$, $c = (0, \ldots, 0, c_p, \ldots, c_{m-p})$ se transforme en 0 par l'homomorphisme (12) revient à dire qu'il existe une chaînette $a = (a_0, \ldots, a_p, 0, \ldots, 0)$ telle que

$$\begin{array}{r} d'a_0 = 0, \\ \delta a_0 + d'a_1 = 0, \\ \cdots\cdots\cdots\cdots \\ \delta a_{p-2} + d'a_{p-1} = 0, \\ \delta a_{p-1} + d'a_p = c_p, \end{array}$$

$$\begin{tikzcd}
0 & & & & & \\
a_0 \arrow[u] \arrow[r] & & & & & \\
& a_1 \arrow[u] \arrow[r] & \cdots & & & \\
& & a_{p-2} \arrow[u] \arrow[r] & & & \\
& & & a_{p-1} \arrow[u] \arrow[r] & c_p & \\
& & & & a_p \arrow[u] &
\end{tikzcd}$$

Mais alors la composante c_p sera nulle pour le cocycle $c' = c - \delta a$, i.e. ce cocycle appartiendra au sous-groupe $Z^{p+1,\, m-p-1}$. Puisque $[c] = [c']$, ceci prouve que le sous-groupe $F^{p+1,\, m-p-1}$ est le noyau de l'homomorphisme (12), et donc, *l'homomorphisme* (13) *est un monomorphisme.*

* * *

Les sous-groupes $F^{p,\, m-p}$ forment une série de sous-groupes emboîtés

$$H^m(C^{\cdot\cdot\cdot}) = F^{0,\, m} \supset F^{1,\, m-1} \supset \ldots \supset F^{m,\, 0} \supset F^{m+1,\, -1} = \{0\} \tag{14}$$

commençant par le groupe $H^m(C^{\cdot\cdot\cdot}) = F^{0,\, m}$ et finissant par le sous-groupe nul $F^{m+1,\, -1} = \{0\}$.

Il est commode d'introduire la terminologie générale appropriée.

Définition 4. On appelle *groupe abélien gradué* une suite arbitraire

$$H^{\cdot} = \{H^0, H^1, \ldots, H^m, \ldots\}$$

de groupes abéliens H^m. On appelle *filtration* d'un groupe gradué $H^\cdot$ une famille $\{F^{p,q}\}$ de groupes abéliens, où $-1 \leqslant p, q < \infty$, telle que

1) chaque groupe $F^{p,q}$ est un sous-groupe du groupe H^{p+q};
2) pour tout $m \geqslant 0$ on a les inclusions

$$H^m = F^{0,m} \supset F^{1,m-1} \supset \ldots \supset F^{m,0} \supset F^{m+1,-1} = \{0\}.$$

On appelle *groupe adjoint* à un groupe gradué $H^\cdot$ à filtration un groupe gradué $\mathrm{Gr}(H^\cdot) = \{\mathrm{Gr}(H^m)\}$ tel que

$$\mathrm{Gr}(H^m) = F^{0,m}/F^{1,m-1} \oplus \ldots \oplus F^{m,0}/F^{m+1,-1}. \tag{15}$$

D'une façon générale, les groupes (15) ne permettent pas encore de restituer les groupes H^m, mais *si le groupe H^m est un espace vectoriel de dimension finie (et les groupes $F^{p,m-p}$ sont ses sous-espaces), le groupe $\mathrm{Gr}(H^m)$ est isomorphe au groupe H^m*. En effet, dans ce cas les deux groupes H^m et $\mathrm{Gr}(H^m)$ sont des espaces vectoriels de même dimension finie (égale à la somme des dimensions des espaces quotients $F^{p,m-p}/F^{p+1,m-p-1}$, $p = 0, \ldots, m$), et nous savons que des espaces vectoriels de même dimension sont isomorphes. [Pour construire l'isomorphisme $H^m \to \mathrm{Gr}(H^m)$, il faut choisir dans chaque sous-espace $F^{p,m-p}$ le sous-espace $R^{p,m-p}$ supplémentaire du sous-espace $F^{p+1,m-p-1}$ (i.e. tel que $F^{p,m-p} = F^{p+1,m-p-1} \oplus R^{p,m-p}$). Dans ces conditions

$$H^m = R^{0,m} \oplus R^{1,m-1} \oplus \ldots \oplus R^{m,0}$$

et les applications de passage au quotient $F^{p,m-p} \to F^{p,m-p}/F^{p+1,m-p-1}$ induisent les isomorphismes $R^{p,m-p} \to F^{p,m-p}/F^{p+1,m-p-1}$ qui composent l'isomorphisme $H^m \to \mathrm{Gr}(H^m)$. (On remarquera que la construction de l'isomorphisme $H^m \to \mathrm{Gr}(H^m)$ comporte un important élément d'arbitraire qu'il est en principe impossible d'éviter.)] □

Remarque 3. La proposition relative à l'isomorphisme des espaces vectoriels H^m et $\mathrm{Gr}(H^m)$ *est valable aussi sans la condition de dimension finie*, puisque le théorème d'existence du sous-espace supplémentaire de tout sous-espace n'implique pas cette condition.

En effet, soit $\mathcal{V}$ un espace vectoriel quelconque. On rappelle (cf. définition 5 de la leçon I.3 et la définition 4 de la leçon I.4) qu'une famille (resp. un ensemble) de vecteurs de $\mathcal{V}$ est *libre* s'il en est de même de toute sous-famille (resp. sous-ensemble) finie, et *complète* si chaque vecteur de $\mathcal{V}$ est une combinaison linéaire des vecteurs d'une sous-famille (resp. sous-ensemble) finie. On appelle *base* une famille de vecteurs libre et complète. Il est clair que la réunion de toute famille totalement ordonnée par inclusion de sous-ensembles libres de $\mathcal{V}$ est libre. Donc, d'après le lemme de Zorn (cf. leçon 10), l'ensemble des sous-ensembles libres de $\mathcal{V}$ possède des éléments maximaux. Chacun de ces éléments étant de toute évidence un sous-ensemble complet, ceci prouve que *chaque espace vectoriel $\mathcal{V}$ possède une base*. De plus, toujours en vertu du lemme de Zorn, chaque sous-ensemble libre de $\mathcal{V}$ est contenu dans une base. En particulier, si $\mathcal{P}$ est

un sous-espace de $\mathcal{V}^c$, toute base $\{x_\alpha\}$ de $\mathcal{P}$ est contenue dans une base $\{x_\alpha, y_\beta\}$ de $\mathcal{V}^c$. L'enveloppe linéaire $\mathfrak{A}$ des vecteurs y_β sera visiblement le sous-espace de $\mathcal{V}^c$ supplémentaire de $\mathcal{P}$.

Notre objectif immédiat est le calcul des groupes adjoints (15) pour le groupe gradué $H^\cdot(C^{\cdot\cdot\cdot}) = \{H^m(C^{\cdot\cdot\cdot})\}$ à filtration (14). Pour cela il nous faut d'abord caractériser l'image de chaque homomorphisme (9) et ensuite décrire le passage des groupes $E_2^{p,q}$ aux groupes $\overline{E}_2^{p,q}$. Ceci fera l'objet de la prochaine leçon.

LEÇON 23

Groupes $E_r^{p,q}$. — Suites spectrales. — Suite spectrale d'un complexe double. — Suite spectrale d'un recouvrement.

Poursuivons l'étude des groupes $E_2^{p,q}$ pour le complexe double $C^{\cdot\cdot} = \{C^{p,q};\ \delta,\ \mathrm{d}\}$.

On sait que (cf. leçon précédente)

a_2) *Un élément* $[c]_2 \in E_2^{p,q}$, $c \in C^{p,q}$, *est défini si et seulement si* $\mathrm{d}'c = 0$ *et existe une cochaîne* c_1 *telle que* $\delta c + \mathrm{d}'c_1 = 0$; *schématiquement*

$$\begin{tikzcd} 0 & \\ c \arrow[u] \arrow[r] & \\ & c_1 \arrow[u] \end{tikzcd} \tag{1}$$

De plus,

b_2) *L'égalité* $[c]_2 = [c']_2$ *est réalisée si et seulement si existent des cochaînes* a *et* b *telles que* $\mathrm{d}'a = 0$ *et* $c' - c = \delta a - \mathrm{d}'b$; *schématiquement*

$$\begin{tikzcd} 0 & & \\ a \arrow[u] \arrow[r] & c' \arrow[r, dash] & c \\ & b \arrow[u] & \end{tikzcd} \tag{2}$$

D'autre part, l'élément $[c]_2$ appartient à l'image de l'homomorphisme $Z^{p,m-p} \to E_2^{p,m-p}$ si et seulement si existe une chaînette (1) susceptible d'être prolongée en la chaînette

$$\begin{tikzcd} 0 & & & & & \\ c \arrow[u] \arrow[r] & & & & & \\ & c_1 \arrow[u] \arrow[r] & & & & \\ & & \ddots & & & \\ & & & c_{q-1} \arrow[u] \arrow[r] & & \\ & & & & c_q \arrow[u] \arrow[r] & 0 \end{tikzcd} \tag{3}$$

$$\begin{aligned} \mathrm{d}'c &= 0, \\ \delta c + \mathrm{d}'c_1 &= 0, \\ &\cdots\cdots\cdots \\ \delta c_{q-1} + \mathrm{d}'c_q &= 0, \\ \delta c_q &= 0, \end{aligned}$$

de longueur $q + 1$, aboutissant à la ligne inférieure du complexe $C^{\cdot\cdot\cdot}$.

Donc, parmi les chaînettes de la forme (1) nous devons choisir celles qui sont prolongeables à une chaînette de la forme (3). Il est naturel de le faire en prolongeant pas à pas la chaînette (1).

Pour prolonger la chaînette (1) d'un maillon, on considère une cochaîne $\delta c_1 \in C^{p+2,\, q-1}$

$$\begin{array}{ccc} 0 & & \\ \uparrow & & \\ c & \rightarrow & \\ & \uparrow & \\ & c_1 & \rightarrow \delta c_1. \end{array}$$

Puisque $d\,\delta c_1 = \delta\, dc_1 = \pm\delta\,\delta c = 0$ et $\delta\,(\delta c_1) = 0$, on a pour la cochaîne δc_1 le diagramme

$$\begin{array}{ccc} 0 & & \\ \uparrow & & \\ \delta c_1 & \rightarrow & \\ & \uparrow & \\ & 0 & \end{array}$$

qui montre que cette cochaîne définit un élément $[\delta c_1]_2 \in E_2^{p+2,\, q-1}$. La cochaîne c_1 de la chaînette (1) pouvant être remplacée par toute cochaîne de la forme $c_1' = c_1 + a$, où $d'a = 0$, la cochaîne δc_1 est définie au terme δa près:

$$\begin{array}{ccc} 0 & & \\ \uparrow & & \\ a & \rightarrow \delta c_1' & - \delta c_1 \\ & & \uparrow \\ & & 0. \end{array}$$

Ceci indique (cf. b_2)) que l'élément $[\delta c_1]_2$ ne dépend pas du choix de la cochaîne c_1 et est défini de façon unique par la cochaîne c. Bien plus, si $[c']_2 = [c]_2$, i.e. (cf. encore b_2)) si

$$\begin{array}{ccc} 0 & & \\ \uparrow & & \\ a & \rightarrow c' & - c \\ & & \uparrow \\ & & b, \end{array}$$

alors

$$\begin{array}{ccc} 0 & & \\ \uparrow & & \\ c' & \rightarrow & \\ & \uparrow & \\ & c_1', & \end{array}$$

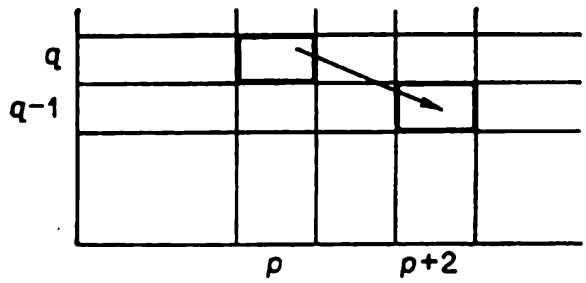
où $c_1' = c_1 + \delta b$, donc $\delta c_1' = \delta c_1$. Par conséquent, $[\delta c_1]_2$ dépend de $[c]_2$ seul, et par suite, la formule

$$d_2 [c]_2 = [\delta c_1]_2$$

définit de façon unique une application

$$d_2 : E_2^{p,q} \to E_2^{p+2,q-1}; \tag{4}$$

schématiquement :



De la définition de l'application d_2 et de la proposition b_2) il s'ensuit immédiatement que

c_2) *L'égalité* $d_2 [a]_2 = [c]_2$ *équivaut à l'existence d'une chaînette de la forme*

$$\begin{array}{ccccc} 0 & & & & \\ \uparrow & & & & \\ a & \to & & & \\ & & \uparrow & & \\ & & a_1 & \to & c \\ & & & & \uparrow \\ & & & & b. \end{array} \tag{5}$$

En particulier $d_2 [c]_2 = 0$ *si et seulement si existe une chaînette de la forme*

$$\begin{array}{ccccc} 0 & & & & \\ \uparrow & & & & \\ c & \to & & & \\ & & \uparrow & & \\ & & c_1 & \to & \\ & & & & \uparrow \\ & & & & c_2. \end{array} \tag{6}$$

Mais si (5) a lieu, (6) aussi aura lieu avec $c_1 = 0$ et $c_2 = 0$. Donc,

$$d_2 \circ d_2 = 0, \tag{7}$$

i.e. pour tous p et q on a l'inclusion

$$B_2^{p,q} \subset Z_2^{p,q},$$

où

$$Z_2^{p,q} = \mathrm{Ker}\,(d_2 : E_2^{p,q} \to E_2^{p+2,q-1})$$

est le noyau de l'application $d_2 : E_2^{p,q} \to E_2^{p+2,q-1}$, et

$$B_2^{p,q} = \mathrm{Im}\,(d_2 : E_2^{p-2,q+1} \to E_2^{p,q}).$$

l'image de l'application $d_2 : E_2^{p-2,q+1} \to E_2^{p,q}$. Posons

$$E_3^{p,q} = Z_2^{p,q}/B_2^{p,q}.$$

Pour l'élément $[c]_2 \in Z_2^{p,q}$ l'élément correspondant du groupe $E_3^{p,q}$ (la classe $[c]_2 + B_2^{p,q}$) sera désigné par le symbole $[c]_3$.

Par définition, l'élément $[c]_3$ est défini si et seulement si l'élément $[c]_2$ l'est et $d_2\,[c]_2 = 0$. D'après la deuxième proposition de c_2) ceci exprime que

a_3) *L'élément* $[c]_3 \in E_3^{p,q}$, $c \in C^{p,q}$, *est défini si et seulement si existe une chaînette de la forme*

$$\begin{tikzcd} 0 & & \\ c \arrow[u] \arrow[r] & & \\ & c_1 \arrow[u] \arrow[r] & \\ & & c_2. \arrow[u] \end{tikzcd} \tag{8}$$

Ceci étant, en vertu de la première proposition de c_2),

b_3) *L'égalité* $[c]_3 = [c']_3$ *est réalisée si et seulement si existe une chaînette de la forme*

$$\begin{tikzcd} 0 & & & \\ a \arrow[u] \arrow[r] & & & \\ & a_1 \arrow[u] \arrow[r] & c' \arrow[r, dash] & c \\ & & b. \arrow[u] & \end{tikzcd}$$

Pour toute chaînette (8) la cochaîne δc_2 est telle que

$$\begin{tikzcd} 0 & \\ \delta c_2 \arrow[u] \arrow[r] & \\ & 0. \arrow[u] \end{tikzcd}$$

Donc, est défini l'élément $[\delta c_2]_2 \in E_2^{p+3,q-2}$, et de plus

$$d_2\,[\delta c_2]_2 = 0.$$

Par suite, est défini l'élément $[\delta c_2]_3$. La différence de deux chaînettes (8) étant de la forme

$$\begin{array}{ccc} 0 & & \\ \uparrow & & \\ a_1 & \rightarrow & \\ & \uparrow & \\ & a_2, & \end{array}$$

un changement de la chaînette (8) se traduit par l'adjonction de l'élément $[\delta a_2]_2 = d_2 [a_1]_2$ à l'élément $[\delta c_2]_2$, donc l'élément $[\delta c_2]_3$ ne change pas. Ce qui exprime que la formule

$$d_3 [c]_2 = [\delta c_2]_3$$

définit de façon unique une application

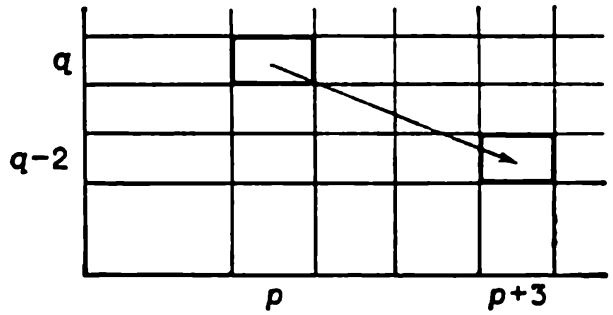
$$d_3 : Z_2^{p,q} \rightarrow E_3^{p+3,q-2}. \tag{9}$$

Pour atteindre une complète analogie avec le cas précédent, on remarquera que si $[c]_2 \in B_2^{p,q}$, i.e. si

$$\begin{array}{ccc} 0 & & \\ \uparrow & & \\ a_1 & \rightarrow & \\ & \uparrow & \\ & a_2 & \rightarrow c, \end{array}$$

il existe alors pour c une chaînette (8) avec $c_1 = 0$ et $c_2 = 0$. Donc, $d_3 [c]_2 = 0$, et par conséquent, l'application (9) induit une application

$$d_3 : E_3^{p,q} \rightarrow E_3^{p+3,q-2}. \tag{10}$$



De la définition de l'application (10) et de la proposition b_3) il s'ensuit immédiatement que

c_3) *L'égalité* $d_3 [a]_3 = [c]_3$ *équivaut à l'existence d'une chaînette de la forme*

$$
(11)\qquad
\begin{array}{ccccccc}
0 & & & & & & \\
\uparrow & & & & & & \\
a & \rightarrow & & & & & \\
 & & \uparrow & & & & \\
 & & a_1 & \rightarrow & & & \\
 & & & & \uparrow & & \\
 & & & & a_2 & \rightarrow & c \\
 & & & & & & \uparrow \\
 & & & & & & b.
\end{array}
$$

En particulier, $d_3\,[c]_3 = 0$ *si et seulement si existe une chaînette de la forme*

$$
(12)\qquad
\begin{array}{ccccccc}
0 & & & & & & \\
\uparrow & & & & & & \\
c & \rightarrow & & & & & \\
 & & \uparrow & & & & \\
 & & c_1 & \rightarrow & & & \\
 & & & & \uparrow & & \\
 & & & & c_2 & \rightarrow & \\
 & & & & & & \uparrow \\
 & & & & & & c_3.
\end{array}
$$

Puisque (11) entraîne l'existence pour la cochaîne c d'une chaînette (12) avec $c_1 = 0$, $c_2 = 0$ et $c_3 = 0$, on voit qu'à l'instar de l'application d_2 l'application d_3 satisfait à la relation

$$d_3 \circ d_3 = 0,$$

et donc sont définis les groupes

$$\overline{E}_4^{p,q} = Z_3^{p,q}/B_2^{p,q},$$

où

$$Z_3^{p,q} = \operatorname{Ker}(d_3 : E_3^{p,q} \rightarrow E_3^{p+3,q-2}),$$

$$B_3^{p,q} = \operatorname{Im}(d_3 : E_3^{p-3,q+2} \rightarrow E_3^{p,q}).$$

En poursuivant cette procédure, on obtient pour tout $r \geqslant 2$ des groupes $E_r^{p,q}$ composés des éléments de la forme $[c]_r$, où $c \in C^{p,q}$, et de plus on aura les propositions suivantes:

a_r) *L'élément* $[c]_r$ *est défini si et seulement si existe une chaînette de la forme*

$$
(13)\qquad
\begin{array}{ccccccc}
0 & & & & & & \\
\uparrow & & & & & & \\
c & \rightarrow & & & & & \\
 & & \uparrow & & & & \\
 & & c_1 & \rightarrow & & & \\
 & & & & \ddots & & \\
 & & & & & \uparrow & \\
 & & & & & c_{r-1}. &
\end{array}
$$

b_r) *L'égalité* $[c]_r = [c']_r$ *est réalisée si et seulement si existe une chaînette de la forme*

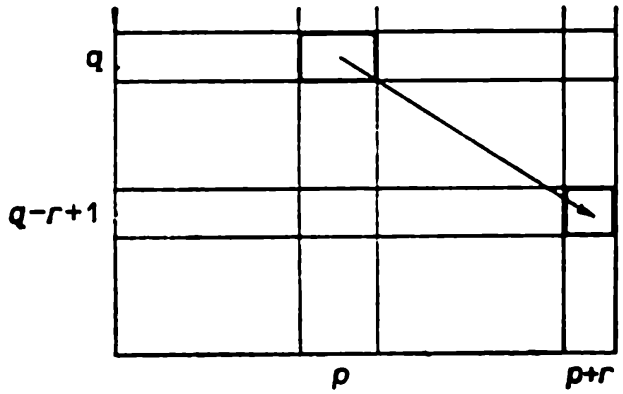
$$\begin{array}{ccccc} 0 & & & & \\ \uparrow & & & & \\ a_1 & \rightarrow & & & \\ & \ddots & & & \\ & & \uparrow & & \\ & & a_{r-2} & \rightarrow & c' - c \\ & & & & \uparrow \\ & & & & b. \end{array} \tag{14}$$

Pour toute chaînette (13), l'élément $[\delta c_{r-1}]_r$ dépend du seul élément $[c]_r$, de sorte que la formule

$$d_r\,[c]_r = [\delta c_{r-1}]_r$$

définit de façon unique une application

$$d_r : E_r^{p,q} \rightarrow E_r^{p+r,\,q-r+1} \tag{15}$$



Ceci étant,

c_r) *L'égalité* $d_r\,[a]_r = [c]_r$ *équivaut à l'existence d'une chaînette de la forme*

$$\begin{array}{ccccc} 0 & & & & \\ \uparrow & & & & \\ a & \rightarrow & & & \\ & \uparrow & & & \\ & a_1 & \rightarrow & & \\ & & \ddots & & \\ & & \uparrow & & \\ & & a_{r-1} & \rightarrow & c \\ & & & & \uparrow \\ & & & & b. \end{array}$$

En particulier, $d_r\,[c]_r = 0$ *si et seulement si existe une chaînette de la forme*

$$\begin{tikzcd}
0 & & & & \\
c \arrow[u] \arrow[r] & & & & \\
& c_1 \arrow[u] \arrow[r] & & & \\
& & \ddots & & \\
& & & c_{r-1} \arrow[u] \arrow[r] & \\
& & & & c_r \arrow[u]
\end{tikzcd}$$

Donc, $d_r \circ d_r = 0$, et par suite, sont définis les groupes

$$E_{r+1}^{p,q} = Z_r^{p,q}/B_r^{p,q},$$

où

$$Z_r^{p,q} = \operatorname{Ker}(d_r : E_r^{p,q} \to E_r^{p+r,q-r+1}),$$

$$B_r^{p,q} = \operatorname{Im}(d_r : E_r^{p-r,q+r-1} \to E_r^{p,q}),$$

etc.

* * *

L'objet construit mérite d'être spécialement défini.

Définition 1. On appelle *suite spectrale* (ou plus précisément *suite spectrale cohomologique du premier quart*) une famille

$$\{E_r^{p,q};\ d_r\} \tag{16}$$

de groupes (ou d'espaces vectoriels) $E_r^{p,q}$ et d'homomorphismes

$$d_r : E_r^{p,q} \to E_r^{p+r,q-r+1}, \tag{17}$$

où $r \geqslant 2$ et $0 \leqslant p, q < \infty$ (du reste, pour éviter toute sorte de restrictions, on conviendra que les groupes $E_r^{p,q}$ sont définis aussi pour les p et $q < 0$ et que de plus $E_r^{p,q} = 0$ si $p < 0$ ou $q < 0$), tels que pour tout $r \geqslant 2$

$$d_r \circ d_r = 0 \tag{18}$$

et

$$E_{r+1}^{p,q} = Z_r^{p,q}/B_r^{p,q}, \tag{19}$$

où

$$Z_r^{p,q} = \operatorname{Ker}(d_r : E_r^{p,q} \to E_r^{p+r,q-r+1}),$$

$$B_r^{p,q} = \operatorname{Im}(d_r : E_r^{p-r,q+r-1} \to E_r^{p,q})$$

(en vertu de (18) on a l'inclusion $B_r^{p,q} \subset Z_r^{p,q}$, et par suite, est défini le groupe quotient (19)).

Les homomorphismes (17) sont généralement appelés *différentielles* de la suite spectrale (16).

Pour tout $r \geqslant 2$ fixe il est commode de ranger les groupes $E_r^{p,q}$ dans une table de la forme

E_r :

$E_r^{0,\,q}$	...	$E_r^{p,\,q}$	...
. . .	. . .	. . .	. . .
$E_r^{0,\,0}$	...	$E_r^{p,\,0}$	...

appelée *r-ième terme* de la suite spectrale (16).

On passe en deux pas de la table E_r à la table E_{r+1}. Au premier pas (« effacement ») on ne laisse dans chaque case (p, q) que les éléments qui se transforment en 0 par la différentielle (17). Ce pas ne modifie pas le contenu de la case (p, q) lorsque la case $(p + r, q - r + 1)$ sur laquelle agit la différentielle (17) ne contient que le zéro. Puisque la case $(p + r, q - r + 1)$ est située à $r - 1$ rangs au-dessous de la case (p, q), l'effacement cesse de toute évidence dans (p, q), et par suite, pour la différentielle d_{q+1}

$$Z_{q+2}^{p,\,q} = Z_{q+3}^{p,\,q} = \ldots$$

Le deuxième pas (« réduction ») consiste à diviser le contenu effacé de la case (p, q) par le sous-groupe des éléments provenant de la case $(p - r, q + r - 1)$ située à r colonnes à gauche. Donc, ce pas ne modifie pas le contenu de la case (p, q) lorsque $p - r < 0$, i.e. pour $r \geqslant p + 1$:

$$B_{p+1}^{p,\,q} = B_{p+2}^{p,\,q} = \ldots$$

Donc, pour tous p et q, le contenu initial $E_2^{p,\,q}$ de la case (p, q) diminue progressivement, lorsque r croît, et finira tôt ou tard par *se stabiliser*, i.e. il existe un $r_0 \geqslant 2$ (plus exactement, $r_0 = \max(p + 1, q + 2)$) tel que

$$E_{r_0}^{p,\,q} = E_{r_0+1}^{p,\,q} = \ldots$$

Posons

$$E_\infty^{p,\,q} = E_{r_0}^{p,\,q}.$$

En vertu de (19), chaque élément y du groupe $E_r^{p,\,q}$ est l'image d'un élément du groupe $E_{r-1}^{p,\,q}$ (appartenant au sous-groupe $Z_{r-1}^{p,\,q}$) qui à son tour est l'image d'un élément du groupe $E_{r-2}^{p,\,q}$ (appartenant à l'image réciproque du sous-groupe $Z_{r-1}^{p,\,q}$ dans $E_{r-2}^{p,\,q}$), etc. Le fait que ce mouvement en arrière nous conduit à un élément x du groupe $E_2^{p,\,q}$ se traduit par la formule $y = [x]_r$. Pour $r = r_0$ on

écrira $[x]_\infty$ au lieu de $[x]_{r_0}$. (Donc, le symbole $[x]_r$, construit plus haut pour les éléments $[c]_r$ à l'aide du complexe double de la suite spectrale, aura la même signification que le symbole $[c]_r$, où c est une cochaîne telle que $x = [c]_2$.)

Si $[x]_r$ est défini pour $x \in E_2^{p,q}$, on dit que x *survit jusqu'à* E_r. Un élément x survit jusqu'à E_{r+1} si et seulement s'il survit jusqu'à E_r, et $d_r [x]_r = 0$.

Il est commode (surtout dans les raisonnements oraux non formels) de considérer que chaque différentielle d_r est une application partielle et multivalente de $E_2^{p,q}$ dans $E_2^{p+r,q-r+1}$. [L'épithète « partielle » signifie qu'en fait d_r est définie seulement sur un sous-groupe du groupe $E_2^{p,q}$ (plus exactement, sur le sous-groupe des éléments survivant jusqu'à E_r), l'épithète « multivalente », que les valeurs de la différentielle d_r appartiennent en fait non pas au groupe $E_2^{p+r,q-r+1}$ mais à un de ses groupes quotients (ou plus exactement à un sous-groupe de ce groupe quotient).] Pour cette raison, l'élément $d_r [x]_r$ est désigné généralement par $d_r x$. Il est défini si et seulement si $d_2 x = 0, \ldots, d_{r-1} x = 0$.

Si $d_r x = 0$, l'élément x s'appelle *cycle de la différentielle* d_r. En particulier, si $d_r x = 0$ pour tout $r \geqslant 2$, l'élément x s'appelle *cycle de toutes les différentielles*. Ceci a lieu si et seulement si est défini l'élément $[x]_\infty$ (l'élément x survit jusqu'à E_∞).

Les cycles des différentielles forment un sous-groupe $Z_\infty^{p,q}$ du groupe $E_2^{p,q}$. Le groupe $E_\infty^{p,q}$ est l'image épimorphe du groupe $Z_\infty^{p,q}$ par l'application

$$x \mapsto [x]_\infty, \quad x \in Z_\infty^{p,q}.$$

Donc, si $B_\infty^{p,q}$ est le noyau de cette application, alors

$$E_\infty^{p,q} = Z_\infty^{p,q}/B_\infty^{p,q}$$

pour tous p et q.

* * *

Retournons maintenant au complexe double $C^{\cdot\cdot} = \{C^{p,q}; \delta, d\}$. Les problèmes posés à la fin de la leçon précédente, savoir la caractérisation de l'image de l'homomorphisme

$$Z^{p,q} \to E_2^{p,q} \tag{20}$$

et du noyau de l'épimorphisme

$$E_2^{p,q} \to \overline{E}_2^{p,q}, \tag{21}$$

se résolvent trivialement en termes de suite spectrale $\{E_r^{p,q}; d_r\}$ construite à l'aide de ce complexe. En effet, par définition, l'élément $x = [c]_2$ du groupe $E_2^{p,q}$ appartient à l'image de l'homomorphisme (20) si et seulement si existe une chaînette de la forme (3) pour la cochaîne c, i.e. (cf. proposition c_r)) si l'élément x est un cycle

des différentielles. Donc, *l'image de l'homomorphisme* (20) *est le sous-groupe* $Z_\infty^{p,q}$.

De façon analogue, le noyau de l'épimorphisme (21) est composé des éléments $[c]_2$ pour lesquels existe une chaînette

$$\begin{array}{ccccccc} 0 & & & & & & \\ \uparrow & & & & & & \\ a_0 & \rightarrow & & & & & \\ & \uparrow & & & & & \\ & a_1 & \rightarrow & & & & \\ & & \ddots & & & & \\ & & & \uparrow & & & \\ & & & a_{p-1} & \rightarrow & c & \\ & & & & & \uparrow & \\ & & & & & b & \end{array}$$

de longueur $p+1$ commençant à la première colonne du complexe $C^{\cdot\cdot}$. Puisque, comme le montre une comparaison directe des définitions, ce sont exactement les éléments du sous-groupe $B_\infty^{p,q}$, ceci prouve que

$$\bar{E}_2^{p,q} = E_2^{p,q}/B_\infty^{p,q}.$$

Donc, $E_\infty^{p,q} \subset \bar{E}_2^{p,q}$, et *le monomorphisme* (13) *de la leçon* 22 *est un isomorphisme du groupe quotient* $F^{p,m-p}/F^{p+1,m-p-1}$ sur $E_\infty^{p,q}$.

Donc, le groupe $\mathrm{Gr}\,(H^m(C^{\cdot\cdot\cdot}))$ adjoint au groupe $H^m(C^{\cdot\cdot\cdot})$ est la somme directe des groupes situés sur la m-ième antidiagonale de la table de E_∞:

$E_\infty^{0,m}$		...			...
	$E_\infty^{1,m-1}$	...			...
. . . . .					
		...	$E_\infty^{m-1,1}$		...
		...		$E_\infty^{m,0}$	...

$$\mathrm{Gr}\,(H\,(C^{\cdot\cdot\cdot})) = E_\infty^{0,m} \oplus E_\infty^{1,m-1} \oplus \ldots \oplus E_\infty^{m-1,1} \oplus E_\infty^{m,0}.$$

Définition 2. Soit $H^{\cdot} = \{H^m\}$ un groupe gradué à filtration

$$H^m = F^{0,m} \supset F^{1,m-1} \supset \ldots \supset F^{m,0} \supset F^{m+1,-1} = \{0\}.$$

On dit qu'une suite spectrale $\{E_r^{p,q};\ \mathrm{d}_r\}$ *converge vers le groupe* $H^{\cdot}$ et on écrit

$$E_r^{p,q} \Rightarrow H^m$$

si

$$E_\infty^{p,q} = F^{p,q}/F^{p+1,q-1}$$

pour tous p et q.

En vertu de cette définition, l'étude que nous venons de faire peut être résumée par le théorème final suivant:

Théorème 1. *Pour tout complexe double* $C^{\cdot\cdot} = \{C^{p,q};\ \delta,\ \mathrm{d}\}$ *il existe une suite spectrale* $\{E_r^{p,q};\ \mathrm{d}_r\}$ *telle que*

$$E_2^{p,q} = H_\delta^p H_\mathrm{d}^q (C^{\cdot\cdot})$$

et

$$E_r^{p,q} \Rightarrow H^m (C^{\cdot\cdot}). \ \square$$

Ce théorème était connu de Leray (mais il semble qu'il ne l'ait pas formulé explicitement).

* * *

Corollaire 1. *Pour tout recouvrement numérotable* $\mathfrak{U}$ *d'une variété différentiable séparée* $\mathscr{X}$ *il existe une suite spectrale*

$$\{E_r^{p,q};\ \mathrm{d}_r\}, \tag{22}$$

telle que

$$E_2^{p,q} = H_\delta^p H_\delta^q (\mathfrak{U}) \tag{23}$$

et

$$E_r^{p,q} \Rightarrow H^m \mathscr{X}. \ \square \tag{24}$$

Nous avons déjà signalé que pour calculer les groupes (23) on pouvait se contenter des renseignements contenus dans les n^{os} a) et b) de la page 320, donc on peut admettre que ces groupes sont connus. Il importe de remarquer qu'*aucune autre information sur la suite spectrale* (22) *ne nous est en principe accessible*. En particulier, dans le cas général on ne peut rien dire sur l'action des différentielles d_r (hormis le fait qu'à partir de la case (p, q) elles atteignent la case $(p + r, q - r + 1)$) ou comment sont les sous-groupes filtrants $F^{p,q}$. Bref, la seule chose que nous sachions et dont nous puissions nous servir est l'existence de la suite spectrale (22) de terme initial (23) et possédant la propriété (24).

L'étonnant c'est que cette information est suffisante dans de nombreuses situations intéressantes et importantes !

Exemple 1 (suite de l'exemple 1 de la leçon 22). Comme déjà indiqué dans l'exemple 1 de la leçon 22, le terme E_2 de la suite spectrale du recouvrement à deux éléments $\{U, V\}$ de la sphère S^n est de la forme

$n-1$		$\mathbb{R}$	...
	. . .	. . .	. . .
0	$\mathbb{R}$		...
	0	1	

Donc, toutes les différentielles d_r agissent soit à partir d'une case, soit sur une case, contenant seulement un zéro. Par conséquent, elles sont toutes nulles et $E_2 = E_\infty$. (Une suite spectrale jouissant de cette propriété est dite *dégénérée*.) Donc, les antidiagonales de la table de E_∞ sont composées uniquement de zéros à l'exception de la 0-ième et de la n-ième qui sont constituées, chacune, d'un groupe $\mathbb{R}$. Par conséquent,

$$H^m S^n = \begin{cases} \mathbb{R} \text{ si } m = 0 \text{ ou } m = n, \\ 0 \text{ dans le cas contraire.} \end{cases}$$

Il s'agit bien évidemment du même calcul qu'à la leçon 20 mais exposé compte tenu des derniers résultats.

Exemple 2. Si $\mathfrak{U}$ est un recouvrement de Leray, alors $E_1^{p,q} = 0$ et *a fortiori* $E_2^{p,q} = 0$ pour $q > 0$. Donc, la suite spectrale du recouvrement dégénère et

$$H^m \mathcal{X} = E_2^{m,0}.$$

D'autre part, il est clair que dans le cas envisagé le complexe $\{E^{\cdot\cdot 0}; \delta^*\}$ n'est autre que le complexe de cochaînes $C^\cdot(\mathfrak{U}; \mathbb{R})$ du recouvrement $\mathfrak{U}$, et par suite,

$$E_2^{m,0} = H^m(\mathfrak{U}; \mathbb{R}) \quad \text{pour tout } m \geqslant 0.$$

Nous retrouvons donc le théorème 1 de la leçon 22 qui s'avère de ce fait être un cas particulier trivial du corollaire 1.

LEÇON 24

Espaces topologiques dénombrables à l'infini et paracompacts. — Variétés paracompactes. — Intégrales dans $\mathbb{R}^n$. — Ensembles cubables et densités sur des variétés arbitraires. — Intégration des densités.

Les méthodes analytiques de calcul des groupes de cohomologie impliquent l'élaboration préliminaire d'un calcul intégral sur les variétés. Commençons cette élaboration par l'exposé des éléments nécessaires de topologie.

On rappelle (cf. définition 5 de la leçon 14) qu'un espace topologique $\mathscr{X}$ est *localement compact* si chacun de ses points possède un voisinage U dont l'adhérence $\bar{U}$ est compacte.

Il est clair que *toute variété séparée $\mathscr{X}$ est localement compacte.* (Ceci n'est pas valable pour les variétés non séparées; par exemple, une droite non séparée présentant un point singulier de multiplicité infinie n est pas un espace localement compact.)

Il est aisé de voir aussi que *si un espace $\mathscr{X}$ satisfait au deuxième axiome de dénombrabilité, est localement compact et séparé, il y existe une base dénombrable $\{U_i\}$ d'éléments U_i dont les adhérences $\bar{U}_i$ sont compactes.* (Pour le prouver, il suffit de choisir des ensembles à adhérences compactes dans une base dénombrable arbitraire de l'espace $\mathscr{X}$. Question de contrôle: pourquoi exige-t-on que l'espace soit séparé?)

Définition 1. On dira qu'un espace topologique $\mathscr{X}$ est *dénombrable à l'infini* s'il est la réunion d'une suite strictement croissante

$$O_1 \subset O_2 \subset \ldots \subset O_n \subset O_{n+1} \subset \ldots$$

d'ensembles ouverts (appelée *exhaustion compacte* de l'espace $\mathscr{X}$) telle que l'adhérence $\bar{O}_n$ de chaque ensemble O_n soit compacte et contenue dans O_{n+1}.

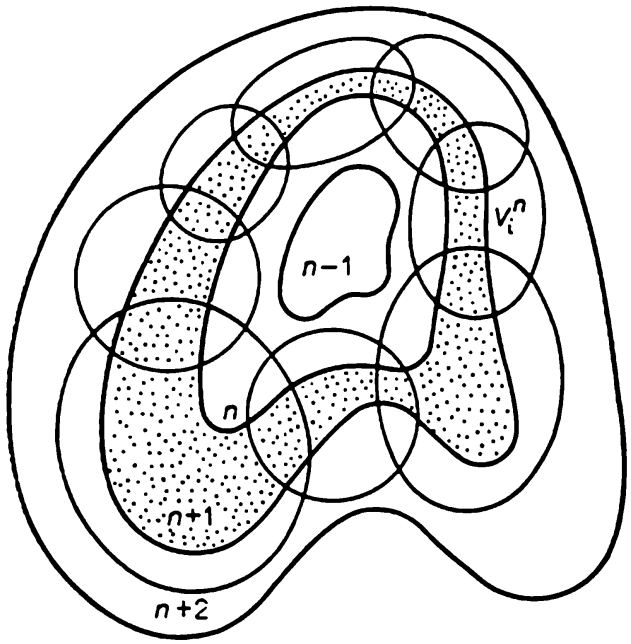
Proposition 1. *Tout espace topologique $\mathscr{X}$ séparé, localement compact et satisfaisant au deuxième axiome de dénombrabilité est dénombrable à l'infini.*

Démonstration. Soit $\{U_i, 1 \leqslant i < \infty\}$ une base dénombrable de l'espace $\mathscr{X}$, composée d'ensembles à adhérences compactes. Posons $O_1 = U_1$ et supposons qu'on ait déjà construit l'ensemble O_{n-1} pour un $n \geqslant 2$. Puisque les ensembles U_i recouvrent $\mathscr{X}$ (et l'ensemble $\bar{O}_{n-1}$ est compact), il existe un $j \geqslant 1$ (qui, pour fixer

les idées, peut être choisi minimal) tel que $\overline{O}_{n-1} \subset U_1 \cup \ldots \cup U_j$. Posons $O_n = U_1 \cup \ldots \cup U_j$. Les ensembles O_n seront ainsi construits par récurrence pour tout $n \geqslant 1$, et il est clair qu'ils forment une exhaustion compacte de $\mathscr{X}$. □

On voit en particulier que *toute variété séparée satisfaisant au deuxième axiome de dénombrabilité est dénombrable à l'infini.*

Définition 2. On dit qu'un recouvrement ouvert $\{U_\alpha\}$ d'un espace topologique $\mathscr{X}$ est *localement fini* si tout point $p \in \mathscr{X}$ possède



un voisinage rencontrant seulement un nombre fini d'éléments de ce recouvrement.

On rappelle (cf. définition 1 de la leçon 8) qu'un recouvrement $\{U_\alpha\}$ de l'espace $\mathscr{X}$ est *inscrit* dans un recouvrement $\{V_\beta\}$ si pour tout α il existe un β tel que $U_\alpha \subset V_\beta$.

Proposition 2. *Si un espace séparé $\mathscr{X}$ est dénombrable à l'infini, dans chacun de ses recouvrements ouverts $\{V_\beta\}$ est inscrit un recouvrement localement fini.*

Démonstration. Soit $\{O_n\}$ une exhaustion compacte arbitraire de $\mathscr{X}$. Pour tout $n \geqslant 1$ les ensembles $V_\beta^n = (O_{n+2} \setminus \overline{O}_{n-1}) \cap V_\beta$ (on conviendra que $O_0 = \varnothing$) forment un recouvrement ouvert de l'ensemble $\overline{O}_{n+1} \setminus O_n$. L'ensemble $\overline{O}_{n+1} \setminus O_n$ étant compact (pourquoi ?), on peut extraire de ce recouvrement un sous-recouvrement fini $\{V_{\beta_1}^n, \ldots, V_{\beta_{m(n)}}^n\}$. Puisque les ensembles $\overline{O}_{n+1} \setminus O_n$ recouvrent $\mathscr{X}$, tous les ensembles de la forme $V_{\beta_i}^n$, $1 \leqslant n < \infty$, $1 \leqslant i \leqslant m(n)$, forment un recouvrement de l'espace $\mathscr{X}$, qui visi-

blement est inscrit dans $\{V_\beta\}$, et puisque aucun ensemble O_k ne rencontre les ensembles $V^n_{\beta_i}$ pour $n > k$ (et donc ne rencontre qu'un nombre fini d'ensembles de la forme $V^n_{\beta_i}$, $n \leqslant k$), le recouvrement $\{V^n_{\beta_i}\}$ est localement fini (pour tout point $p \in \mathscr{X}$ il existe un k tel que $p \in O_k$, donc l'ensemble O_k est un voisinage de ce point rencontrant seulement un nombre fini d'éléments du recouvrement $\{V^n_{\beta_i}\}$). □

Remarque 1. En topologie générale on dit qu'un espace topologique $\mathscr{X}$ est *paracompact* si l'on peut inscrire un recouvrement localement fini dans chacun de ses recouvrements ouverts. Dans cette terminologie la proposition 2 exprime que *tout espace séparé dénombrable à l'infini est paracompact.*

* * *

La proposition 2 peut être précisée lorsque $\mathscr{X}$ est une variété différentiable.

On dira qu'un recouvrement ouvert $\{U_\alpha\}$ d'une variété différentiable $\mathscr{X}$ est *proprement numérotable* s'il existe une partition de l'unité $\{\eta_\alpha\}$, subordonnée à ce recouvrement, telle que $\eta_\alpha = 0$ à l'extérieur d'un ensemble compact $C_\alpha \subset \mathscr{X}$.

Proposition 3. *Dans tout recouvrement ouvert $\{V_\beta\}$ d'une variété différentiable séparée $\mathscr{X}$ satisfaisant au deuxième axiome de dénombrabilité, on peut inscrire un recouvrement $\{U_\alpha\}$ localement fini proprement numérotable.*

Démonstration. La variété $\mathscr{X}$ est justiciable de la proposition 2, car dénombrable à l'infini en vertu de la proposition 1. Soit $\{V^n_{\beta_i}\}$ le recouvrement construit dans la démonstration de la proposition 2. Tout point $p \in V^n_{\beta_i}$ possède un voisinage de coordonnées $U_{p,\,i,\,n}$ à adhérence compacte, contenu dans $V^n_{\beta_i}$. En vertu de la proposition 2 de la leçon 14, il existe des ensembles ouverts $V_{p,\,i,\,n}$ et $W_{p,\,i,\,n}$ tels que

$$p \in W_{p,\,i,\,n},\ \overline{W}_{p,\,i,\,n} \subset V_{p,\,i,\,n},\ \overline{V}_{p,\,i,\,n} \subset U_{p,\,i,\,n},$$

et le couple $(V_{p,\,i,\,n},\ W_{p,\,i,\,n})$ possède une fonction d'Urysohn. Pour tout i fixe, les ensembles $W_{p,\,i,\,n}$ forment un recouvrement fini de l'ensemble compact $\bar{O}_{n+1} \setminus O_n$. Extrayons un sous-recouvrement fini de ce recouvrement, effectuons cette procédure pour tous les i et désignons par W_α les ensembles $W_{p,\,i,\,n}$ obtenus. Soient V_α et U_α respectivement les ensembles $V_{p,\,i,\,n}$ et $U_{p,\,i,\,n}$.

Par construction

a) pour tout α on a les inclusions

$$\overline{W}_\alpha \subset V_\alpha,\quad \overline{V}_\alpha \subset U_\alpha,$$

et de plus l'ensemble $\overline{U}_\alpha$ (donc $\overline{W}_\alpha$ et $\overline{V}_\alpha$) est compact ;

b) pour chaque couple (V_α, W_α) il existe une fonction d'Urysohn

$$\varphi_\alpha : \mathscr{X} \to \mathbb{R}$$

(qui rappelons-le est telle que $\varphi_\alpha = 1$ sur $\overline{W}_\alpha$, et $\varphi_\alpha = 0$ à l'extérieur de V_α);

c) la famille $\{W_\alpha\}$, donc chacune des familles $\{V_\alpha\}$ et $\{U_\alpha\}$ est un recouvrement ouvert de l'espace $\mathscr{X}$ (qui visiblement est localement fini et inscrit dans le recouvrement $\{V_\beta\}$).

Le recouvrement $\{V_\alpha\}$ étant localement fini, il en est de même de la famille de fonctions $\{\varphi_\alpha\}$. Donc, est définie la fonction

$$\varphi = \sum_\alpha \varphi_\alpha.$$

Puisque $\varphi_\alpha = 1$ sur $\overline{W}_\alpha$, et $\varphi_\alpha \geqslant 0$ sur $\mathscr{X}$, on a $\varphi \geqslant 1$ sur $\mathscr{X}$ et en particulier $\varphi \neq 0$ sur $\mathscr{X}$. Donc, sont définies les fonctions

$$\eta_\alpha = \frac{\varphi_\alpha}{\varphi}.$$

Les fonctions η_α forment une partition de l'unité, subordonnée de toute évidence au recouvrement $\{U_\alpha\}$. Comme $\eta_\alpha = 0$ à l'extérieur de l'ensemble compact $\overline{V}_\alpha$, ceci prouve que le recouvrement $\{U_\alpha\}$ est proprement numérotable. Ce recouvrement étant localement fini et inscrit dans le recouvrement $\{V_\beta\}$, ceci prouve entièrement la proposition 3. □

Corollaire 1. *Chaque recouvrement ouvert $\{V_\beta\}$ d'une variété différentiable séparée $\mathscr{X}$ satisfaisant au deuxième axiome de dénombrabilité est numérotable.*

D é m o n s t r a t i o n. En vertu de la proposition 3 il suffit de montrer que *le recouvrement $\{V_\beta\}$ est numérotable si un recouvrement $\{U_\alpha\}$ numérotable y est inscrit.* Choisissons pour tout indice α un indice $\beta(\alpha)$ tel que $U_\alpha \subset V_{\beta(\alpha)}$ et considérons une partition de l'unité $\{\eta_\alpha\}$, subordonnée au recouvrement $\{U_\alpha\}$. Soit

$$\zeta_\beta = \sum_{\beta(\alpha)=\beta} \eta_\alpha,$$

où la sommation est étendue à tous les α tels que $\beta(\alpha) = \beta$. (Si de tels α n'existent pas, alors $\zeta_\beta = 0$.) Il est clair que la fonction ζ_β est définie et positive. De plus, puisque $\zeta_\beta(p) \neq 0$ lorsqu'il existe un α tel que $\beta = \beta(\alpha)$ et $\eta_\alpha(p) \neq 0$, la famille de fonctions $\{\zeta_\beta\}$ est localement finie, et comme chaque fonction η_α figure additivement dans une fonction ζ_β et une seule, il vient

$$\sum_\beta \zeta_\beta = \sum_\alpha \eta_\alpha = 1.$$

Enfin, si $p \notin V_\beta$, alors $p \in U_\alpha$ pour tout α tel que $\beta(\alpha) = \beta$, donc $\eta_\alpha(p) = 0$. Par conséquent, $\zeta_\beta(p) = 0$. Donc, $\{\zeta_\beta\}$ est une

partition de l'unité, subordonnée au recouvrement $\{V_\beta\}$. Ceci prouve que ce recouvrement est numérotable. □

Le corollaire 1 exprime que *toute variété différentiable séparée satisfaisant au deuxième axiome de dénombrabilité est paracompacte.* Cette assertion traduit le contenu du théorème 2 de la leçon 22 que l'on peut dorénavant considérer comme démontré.

Remarque 2. On démontre — un théorème difficile ! — que *chaque composante d'un espace métrisable et localement compact satisfait au deuxième axiome de dénombrabilité.* D'autre part, on montrera dans le prochain ouvrage que *toute variété différentiable séparée et paracompacte est métrisable.* Donc, *dans la classe des variétés différentiables séparées et connexes la paracompacité équivaut au deuxième axiome de dénombrabilité.*

Remarque 3. En vertu de la proposition 3, chaque variété différentiable séparée satisfaisant au deuxième axiome de dénombrabilité est paracompacte en tant qu'espace topologique (cf. remarque 1 ci-dessus). Etant donné qu'un espace topologique est paracompact si et seulement s'il en est de même de chacune de ses composantes, on en déduit immédiatement, d'après ce qui a été dit dans la remarque 2, qu'une variété paracompacte séparée est paracompacte aussi en tant qu'espace topologique. D'autre part, en compliquant légèrement la démonstration de la proposition 3, on démontre que la conclusion de cette proposition (donc, du corollaire 1) reste valable pour toute variété différentiable séparée qui est paracompacte en tant qu'espace topologique. Donc, *une variété séparée est paracompacte si et seulement si elle l'est en tant qu'espace topologique.*

Remarque 4. Il est évident que la notion de recouvrement numérotable se transpose immédiatement à tout espace topologique (il suffit de remplacer les fonctions différentiables par des fonctions continues dans la définition de la partition de l'unité), et il est aisé de voir que *si tout recouvrement ouvert d'un espace topologique $\mathscr{X}$ est numérotable, cet espace est paracompact* (puisque pour chaque partition de l'unité $\{\eta_\alpha\}$ les ensembles $O_\alpha = \{p \in \mathscr{X};\ \eta_\alpha(p) \neq 0\}$ forment un recouvrement ouvert, localement fini, inscrit dans tout recouvrement auquel est subordonnée la partition $\{\eta_\alpha\}$). Il est intéressant de noter que la réciproque est vraie dans la classe des espaces séparés, i.e. *tout recouvrement ouvert d'un espace séparé paracompact est numérotable.* [Remarquons que la proposition de la remarque 3 ne découle pas encore directement de là pour les variétés différentiables, car les recouvrements ouverts numérotables d'une variété différentiable ne constituent à priori qu'une partie de ses recouvrements ouverts numérotables en tant qu'espace topologique.] La démonstration de cette assertion est loin d'être simple et fait intervenir une foule de théorèmes compliqués de topologie générale. Elle implique en premier lieu le théorème d'Urysohn qui affirme que dans un espace normal séparé, pour tout couple (V, W) d'ensembles

ouverts satisfaisant à la relation $\overline{W} \subset V$, il existe une fonction continue égale à l'unité sur W et à zéro à l'extérieur de V, et ensuite la proposition 6 de la leçon 9 relative à l'existence des applications coercives (qui rappelons-le n'a été que partiellement démontrée). A l'aide de ces deux théorèmes on peut déjà démontrer sans grande peine que *tout recouvrement ouvert localement fini d'un espace normal séparé est numérotable*. Après quoi il ne reste plus qu'à prouver le théorème (assez compliqué aussi!) de Dieudonné qui dit que *tout espace paracompact séparé est normal*.

* * *

Nous pouvons passer maintenant à l'élaboration d'un calcul intégral sur les variétés différentiables.

On sait du cours d'analyse qu'un ensemble A de l'espace $\mathbb{R}^n$ est un *ensemble négligeable au sens de Jordan* (ou un *ensemble de mesure* 0) si pour tout $\varepsilon > 0$ il existe une famille finie de boules (ou, ce qui est équivalent, de cubes ou de parallélépipèdes) de l'espace $\mathbb{R}^n$ recouvrant A, de volume total $< \varepsilon$. (Comparer avec la définition d'un ensemble négligeable au sens de Lebesgue dans la leçon 14, où l'on suppose l'existence d'une famille dénombrable.)

On dit qu'un sous-ensemble $\overline{D} \subset \mathbb{R}^n$ est *cubable* (ou *mesurable au sens de Jordan*) s'il est borné (i.e. son adhérence $\overline{D}$ est compacte) et sa frontière $\operatorname{Fr} D$ est un ensemble négligeable. (On appelle frontière $\operatorname{Fr} D$ d'un sous-ensemble D d'un espace topologique l'ensemble $\overline{D} \setminus \operatorname{Int} D$.)

Pour éviter les restrictions, on admettra que toutes les fonctions envisagées sont définies sur l'espace $\mathbb{R}^n$ *tout entier*. Cette condition ne restreint pas la généralité, car toute fonction f définie sur un sous-ensemble D de $\mathbb{R}^n$ peut être prolongée à 0 à l'extérieur de D sans que l'intégrale change, c'est-à-dire qu'on peut admettre que $f(\mathbf{x}) = 0$ pour $\mathbf{x} \notin D$.

On dit qu'une fonction $f\colon \mathbb{R}^n \to \mathbb{R}$ est *à support borné* s'il existe un ensemble compact $C \subset \mathbb{R}^n$ tel que $f = 0$ à l'extérieur de C; *localement bornée* si elle est bornée sur tout ensemble compact $C \subset \mathbb{R}^n$; *presque continue* si elle est continue sur tout ensemble compact $C \subset \mathbb{R}^n$ à l'extérieur d'un ensemble négligeable, i.e. s'il existe une fonction f_C continue sur C et un ensemble négligeable $A \subset C$ tel que $f = f_C$ sur $C \setminus A$.

On sait du cours d'analyse que *toute fonction f presque continue et localement bornée est intégrable sur chaque ensemble cubable D*, i.e. il existe une intégrale finie

$$\int_D f(\mathbf{x})\, d\mathbf{x}. \tag{1}$$

Ceci étant, l'intégrale (1) ne change pas si l'on modifie arbitrairement la fonction f sur un ensemble négligeable ou si l'on ajoute ou retranche un ensemble négligeable de l'ensemble D.

[Nous avons en vue ici une intégrale au sens de Riemann. Si l'on comprend l'intégrale (1) au sens de Lebesgue, au lieu de la cubabilité on peut exiger la mesurabilité de D. Mais nous n'aurons pas besoin de cette généralisation.]

Lorsque la fonction f est à support borné, l'intégrale (1) prend la même valeur sur tout ensemble cubable D à l'extérieur duquel f est nulle. Cette valeur

commune sera appelée *intégrale de f sur* $\mathbf{R}^n$ et désignée par le symbole

$$\int_{\mathbf{R}^n} f(\mathbf{x})\,d\mathbf{x}$$

ou par le symbole

$$\int f(\mathbf{x})\,d\mathbf{x}. \tag{2}$$

Puisque pour tout ensemble cubable D et toute fonction f intégrable sur D on a

$$\int_D f(\mathbf{x})\,d\mathbf{x} = \int_{\mathbf{R}^n} (f\chi_D)(\mathbf{x})\,d\mathbf{x},$$

où χ_D est la *fonction caractéristique* de l'ensemble D, i.e.

$$\chi_D(\mathbf{x}) = \begin{cases} 1 & \text{si} \quad \mathbf{x} \in D, \\ 0 & \text{si} \quad \mathbf{x} \notin D, \end{cases}$$

on peut au besoin se contenter seulement de l'intégrale (2).

Si U et V sont des sous-ensembles ouverts de l'espace $\mathbf{R}^n$, pour toute application différentiable $\varphi\colon U \to V$ et tout ensemble $A \subset U$ négligeable, l'ensemble φA est négligeable aussi. [La démonstration de ce fait pour un ensemble négligeable au sens de Lebesgue, produite dans la leçon 14, s'étend *ad litteram* à un ensemble négligeable au sens de Jordan.] Il s'ensuit que pour tout ensemble cubable $D \subset U$ tel que $\overline{D} \subset U$ l'ensemble $\varphi D \subset V$ est cubable. En particulier, ceci est vrai lorsque φ est un difféomorphisme. Bien plus, on démontre en analyse que dans ce cas, pour toute fonction f intégrable sur l'ensemble φD, la fonction $(f \circ \varphi)\,|\,J_\varphi\,|$, où

$$J_\varphi = \det \left\| \frac{\partial \varphi^i}{\partial x^j} \right\|, \quad i, j = 1, \ldots, n, \tag{3}$$

est le jacobien du difféomorphisme φ, est intégrable sur D et l'on a l'égalité

$$\int_{\varphi D} f(\mathbf{x})\,d\mathbf{x} = \int_D (f \circ \varphi)\,|J_\varphi|\,(\mathbf{x})\,d\mathbf{x}. \tag{4}$$

Cette proposition est connue sous le nom de théorème de changement des variables. Pour les intégrales (2) la formule (4) s'écrit

$$\int f(\mathbf{x})\,d\mathbf{x} = \int (f \circ \varphi)\,|J_\varphi|\,(\mathbf{x})\,d\mathbf{x}, \tag{5}$$

où f est une fonction intégrable à support borné, nulle à l'extérieur de V, et J_φ le jacobien (3) prolongé à 0 à l'extérieur de U comme convenu ci-dessus.

* * *

Pour bâtir une théorie de l'intégration sur une variété différentiable $\mathscr{X}$ (qui sera supposée séparée et paracompacte) il nous faut commencer par la définition d'un sous-ensemble cubable sur $\mathscr{X}$.

Définition 3. Un sous-ensemble A d'une variété différentiable $\mathscr{X}$ à n dimensions s'appelle *ensemble négligeable* s'il existe une

famille finie de cartes $(U_1, h_1), \ldots, (U_N, h_N)$ telle que $A \subset \bigcup U_i$ et pour tout $i = 1, \ldots, N$, l'ensemble $h_i(A \cap U_i)$ est négligeable dans $\mathbb{R}^n$ (cf. définition 6 de la leçon 14). Un sous-ensemble $D \subset \mathscr{X}$ est dit *cubable* si son adhérence $\overline{D}$ est compacte et sa frontière

$$\operatorname{Fr} D = \overline{D} \setminus \operatorname{Int} D$$

est un ensemble négligeable.

Exercice 1. Montrer que A est un ensemble négligeable si et seulement si pour toute famille finie de cartes $(U_1, h_1), \ldots, (U_N, h_N)$ recouvrant A (i.e. telle que $A \subset \bigcup U_i$) les ensembles $h_i(A \cap U_i)$, $1 \leqslant i \leqslant N$, sont des ensembles négligeables dans $\mathbb{R}^n$.

La situation est plus délicate avec les expressions sous le signe d'intégration. Pour qu'une intégrale ne dépende pas du choix des coordonnées locales, il faut que l'intégrant se transforme par un changement de coordonnées conformément aux formules (4) ou (5). Ceci nous conduit à la définition suivante:

Définition 4. Supposons que dans chaque carte (U, h) d'une variété $\mathscr{X}$ est définie une fonction $\rho^{(U, h)}: U \to \mathbb{R}$ (qui dans la suite sera brièvement désignée par ρ^U) et que pour deux cartes sécantes quelconques (U, h) et (V, k) on a l'égalité

$$\rho^V = \rho^U \left| \det \frac{\partial k}{\partial h} \right| \quad \text{sur } U \cap V, \tag{6}$$

où $\det \frac{\partial k}{\partial h}$ est le jacobien du difféomorphisme $\varphi = k \circ h^{-1}$ (traité comme une fonction sur $U \cap V \subset U$, i.e. lié à son jacobien J_φ dans $\mathbb{R}^n$ par la formule $\det \frac{\partial k}{\partial h} = J_\varphi \circ h$). La famille $\{\rho^U\}$ est alors appelée *densité* sur $\mathscr{X}$.

Exemple 1. Une forme différentielle ω de degré n sur une variété $\mathscr{X}$ (où comme toujours $n = \dim \mathscr{X}$) se représente dans chaque carte $(U, h) = (U, x^1, \ldots, x^n)$ par

$$\omega = w^U \, dx^1 \wedge \ldots \wedge dx^n,$$

où w^U est une fonction différentiable sur U. Ceci étant, pour deux cartes quelconques (U, h) et (V, k) les fonctions w^U et w^V correspondantes sont liées par la formule

$$w^V = w^U \det \frac{\partial k}{\partial h} \quad \text{sur} \quad U \cap V.$$

Donc, en posant $\rho^U = |w^U|$, on définit une densité sur $\mathscr{X}$.

Une densité $\rho = \{\rho^U\}$ est dite *différentiable*, *localement bornée* et *presque continue* si les fonctions ρ^U le sont respectivement.

Les densités localement bornées et presque continues sur $\mathscr{X}$ forment un espace vectoriel $\Pi\mathscr{X}$ qui est naturellement un module sur l'anneau des fonctions localement bornées et presque continues

sur $\mathscr{X}$. De façon analogue, l'espace vectoriel $\Pi_{\mathrm{dif}}\,\mathscr{X}$ des densités différentiables est un module sur l'anneau des fonctions différentiables.

On dira qu'une densité ρ est *nulle* (resp. *strictement positive*) en un point $p \in \mathscr{X}$ si pour tout voisinage de coordonnées U contenant le point p on a l'égalité $\rho^U(p) = 0$ (resp. l'inégalité $\rho^U(p) > 0$). La fonction $\det \frac{\partial k}{\partial h}$ ne s'annulant en aucun point, on déduit de la relation (6) que cette définition est intrinsèque (ne dépend pas du choix du voisinage de coordonnées U).

[On remarquera qu'il est absurde de parler de la valeur de la densité ρ en un point p.]

On dit qu'une densité sur une variété $\mathscr{X}$ est *à support borné* s'il existe un ensemble compact $C \subset \mathscr{X}$ tel que $\rho = 0$ à l'extérieur de C. Les densités à support borné forment un sous-espace (un sous-module) $\Pi_{\mathrm{s.b}}\,\mathscr{X}$ de l'espace (du module) $\Pi\mathscr{X}$. Toute densité est à support borné sur (et seulement sur!) une variété compacte.

Une densité strictement positive en chaque point $p \in \mathscr{X}$ s'appelle *densité volumique* sur $\mathscr{X}$. La densité volumique est traditionnellement désignée par $\mathrm{d}v$. Pour chaque densité ρ la formule

$$f = \frac{\rho^U}{\mathrm{d}v^U} \text{ sur } U$$

(pour une densité volumique donnée $\mathrm{d}v$) définit de façon unique sur $\mathscr{X}$ une fonction f telle que $\rho = f\,\mathrm{d}v$ sur $\mathscr{X}$. Donc, *si sur une variété $\mathscr{X}$ il existe une densité volumique différentiable, les espaces vectoriels $\Pi\mathscr{X}$ et $\Pi_{\mathrm{dif}}\,\mathscr{X}$ sont des modules libres à une dimension (le premier sur l'anneau des fonctions localement bornées et presque continues, le second sur l'anneau des fonctions différentiables)*. [Leurs sous-modules composés des densités à support borné ne seront généralement ni libres ni à une dimension.]

Pour une densité volumique donnée l'intégrale

$$\int_D \mathrm{d}v$$

s'appelle *volume* de l'ensemble cubable D.

Exemple 2. Soit $\mathscr{X}$ une sous-variété à deux dimensions d'un espace euclidien (une surface). Alors (cf. leçon 3) dans chaque voisinage de coordonnées $U \subset \mathscr{X}$ (qui est une surface élémentaire au sens de la définition 2 de la leçon 3) sont définies des fonctions E, F et G (les coefficients de la première forme quadratique) telles que le déterminant

$$\begin{vmatrix} E & F \\ F & G \end{vmatrix} = EG - F^2 > 0$$

est multiplié, dans tout changement de coordonnées, par le carré du déterminant de la matrice de passage. Donc, la formule

$$\rho^U = \sqrt{EG - F^2}$$

définit sur $\mathscr{X}$ une densité volumique (ou plus précisément une densité surfacique). Cette densité est désignée par $d\sigma$ (ou par un symbole semblable, par exemple dS) et appelée *élément d'aire* de la surface $\mathscr{X}$. (Cf. leçon 3.)

Exercice 2. Montrer que sur toute variété différentiable paracompacte séparée $\mathscr{X}$ il existe une densité volumique différentiable dv.

Donc, sur une variété paracompacte séparée, on peut identifier les densités aux fonctions en choisissant une certaine densité volumique.

* * *

Nous pouvons maintenant formuler le théorème fondamental d'existence et d'unicité de l'intégrale. Nous le ferons même en deux versions.

Soit $\mathscr{X}$ une variété différentiable paracompacte séparée.

Théorème 1. *Pour tout ensemble cubable $D \subset \mathscr{X}$ il existe une seule fonctionnelle linéaire*

$$\int_D : \Pi\mathscr{X} \to \mathbb{R} \tag{7}$$

possédant les deux propriétés suivantes:

a) *Si $D = D_1 \sqcup D_2$ (i.e. si $D = D_1 \cup D_2$ et $D_1 \cap D_2 = \varnothing$), alors*

$$\int_D \rho = \int_{D_1} \rho + \int_{D_2} \rho$$

pour toute densité $\rho \in \Pi\mathscr{X}$.

b) *Si pour une densité $\rho \in \Pi\mathscr{X}$ il existe une carte (U, h) telle que $\rho = 0$ à l'extérieur de U, alors*

$$\int_D \rho = \int_{h(D \cap U)} (\rho^U \circ h^{-1})(\mathbf{x})\, d\mathbf{x},$$

où l'intégrale de droite est une intégrale dans $\mathbb{R}^n$.

Théorème 1a. *Il existe une seule fonctionnelle linéaire*

$$\int : \Pi_{s.b.}\mathscr{X} \to \mathbb{R} \tag{8}$$

telle que

$$\int \rho = \int_{h(U)} (\rho^U \circ h^{-1})(\mathbf{x})\, d\mathbf{x} \tag{9}$$

pour toute carte (U, h) *et toute densité* ρ *à support borné nulle à l'extérieur de* U.

La fonctionnelle (8) est désignée aussi par le symbole $\int_{\mathscr{X}}$.

Pour déduire le théorème 1a du théorème 1, il suffit pour toute densité $\rho \in \Pi_{s.b}\,\mathscr{X}$ de poser

$$\int \rho = \int_D \rho, \tag{10}$$

où D est un ensemble cubable à l'extérieur duquel la densité ρ est nulle. (Il est clair qu'un tel ensemble existe et que l'intégrale (10) est indépendante de son choix.) Réciproquement, puisque pour toute densité $\rho \in \Pi\mathscr{X}$ et tout ensemble cubable $D \subset \mathscr{X}$ la densité $\chi_D\rho$, où χ_D est la fonction caractéristique de D, est de toute évidence une densité à support borné, localement bornée et presque continue, on peut construire l'intégrale (7) à l'aide de l'intégrale (8) en posant

$$\int_D \rho = \int \chi_D\rho.$$

(La propriété b) est évidente pour l'intégrale (7), quant à la propriété a), elle résulte du fait que $\chi_{D_1 \cup D_2} = \chi_{D_1} + \chi_{D_2}$.) Donc, les théorèmes 1 et 1a résultent l'un de l'autre et l'on ne prouvera par conséquent que l'un d'eux, par exemple le second.

Démonstration du théorème 1a. Commençons comme toujours par prouver l'unicité.

La variété $\mathscr{X}$ possède un recouvrement localement fini numérotable $\{U_\alpha\}$ composé de voisinages de coordonnées, car elle est séparée et paracompacte. Soit $\{\eta_\alpha\}$ une partition de l'unité subordonnée au recouvrement $\{U_\alpha\}$. Pour toute densité $\rho \in \Pi_{s.b}\,\mathscr{X}$, seul un nombre fini de densités $\eta_\alpha\rho$ seront non nulles (prouvez-le !), donc on aura la formule

$$\rho = \sum_\alpha \eta_\alpha\rho.$$

Comme $\eta_\alpha\rho = 0$ à l'extérieur de U_α, on en déduit (cf. formule (9)) que

$$\int \rho = \sum_\alpha \int_{h_\alpha(U_\alpha)} (\eta_\alpha\rho^{U_\alpha} \circ h_\alpha^{-1})(\mathbf{x})\,d\mathbf{x}, \tag{11}$$

où h_α est une application de coordonnées $U_\alpha \to \mathbb{R}^n$. Ceci prouve l'unicité, puisque le second membre de la formule (11) est indépendant du choix de la fonctionnelle (8).

Pour prouver l'existence, on définira la fonctionnelle (8) par la formule (11). Il est clair que cette fonctionnelle est linéaire. De plus, si $\rho = 0$ à l'extérieur de U, le théorème de changement des

variables nous dit que (cf. formule (4))

$$\int_{h_\alpha(U_\alpha)} (\eta_\alpha \rho^{U_\alpha} \circ h_\alpha^{-1})(\mathbf{x})\,d\mathbf{x} = \int_{h_\alpha(U_\alpha \cap U)} (\eta_\alpha \rho^{U_\alpha} \circ h_\alpha^{-1})(\mathbf{x})\,d\mathbf{x} =$$

$$= \int_{h(U_\alpha \cap U)} (h_\alpha \rho^{U_\alpha} \circ h^{-1}) \mid J_\varphi \mid (\mathbf{x})\,d\mathbf{x},$$

où $\varphi = h \circ h_\alpha^{-1}$. Puisque par définition

$$J_\varphi = \det \frac{\partial h}{\partial h_\alpha} \circ h^{-1}$$

et

$$(\eta_\alpha \rho^{U_\alpha} \circ h^{-1}) \cdot \left| \det \frac{\partial h}{\partial h_\alpha} \circ h^{-1} \right| = \eta_\alpha \rho^U \circ h^{-1}$$

(cf. formule (6)), on a

$$\int_{h_\alpha(U_\alpha)} (\eta_\alpha \rho^{U_\alpha} \circ h_\alpha^{-1})(\mathbf{x})\,d\mathbf{x} = \int_{h(U_\alpha \cap U)} (\eta_\alpha \rho^{U} \circ h^{-1})(\mathbf{x})\,d\mathbf{x} =$$

$$= \int_{h(U)} (\eta_\alpha \rho^{U} \circ h^{-1})(\mathbf{x})\,d\mathbf{x}$$

et, par suite,

$$\int \rho = \sum_\alpha \int_{h(U)} (\eta_\alpha \rho^U \circ h^{-1})(\mathbf{x})\,d\mathbf{x} =$$

$$= \Big(\sum_\alpha \eta_\alpha\Big) \int_{h(U)} (\rho^U \circ h^{-1})(\mathbf{x})\,d\mathbf{x} = \int_{h(U)} (\rho^U \circ h^{-1})(\mathbf{x})\,d\mathbf{x}. \quad \square$$

Si une densité volumique dv est donnée sur $\mathscr{X}$, pour toute fonction f à support borné, localement bornée et presque continue sur $\mathscr{X}$, est définie l'intégrale

$$\int f\,dv = \int_{\mathscr{X}} f\,dv$$

appelée *intégrale de f sur $\mathscr{X}$ par rapport à* dv.

Exemple 3 (suite de l'exemple 2). Pour toute fonction f à support borné (et, disons, différentiable) est définie sur une surface $\mathscr{X}$ de l'espace euclidien l'intégrale

$$\int_{\mathscr{X}} f\,d\sigma, \tag{12}$$

où $d\sigma$ est l'élément d'aire sur $\mathscr{X}$ (cf. exemple 2 plus haut). Les intégrales de la forme (12) s'appellent *intégrales de surface de première*

espèce. En analyse on démontre qu'elles sont les limites de sommes intégrales obtenues de façon naturelle en subdivisant la surface en un nombre (qui tend vers l'infini) de surfaces élémentaires.

[Cf. les remarques faites à la leçon 3 sur la mesure des aires sur une surface.]

Exemple 4. Soit $\gamma : [a, b] \to \mathbb{R}^n$ une courbe différentiable dans l'espace euclidien $\mathbb{R}^n$, qui n'est assujettie en général à aucune condition de régularité, ou en d'autres termes une fonction vectorielle différentiable

$$\mathbf{r} = \mathbf{r}(t), \quad a \leqslant t \leqslant b$$

(cf. leçon 1). Pour toute fonction $f = f(\mathbf{r})$ sur $\mathbb{R}^n$, dont le domaine de définition contient le support $\gamma\,[a, b]$ de la courbe, posons

$$\int_\gamma f\,\mathrm{d}s = \int_a^b f(\mathbf{r}(t))\,|\,\mathbf{r}'(t)\,|\,\mathrm{d}t. \tag{13}$$

Cette intégrale s'appelle *intégrale curviligne de première espèce* de la fonction f le long de la courbe γ. C'est la limite de sommes intégrales de la forme $\sum f(\mathbf{r}_i)\,s_i$, où s_i sont les longueurs des segments d'une subdivision arbitraire de γ, $\mathbf{r}_i$ des points de ces segments. D'après les définitions générales introduites plus haut, cette intégrale est une intégrale de la fonction $f \circ \gamma : t \mapsto f(\mathbf{r}(t))$ étendue à la variété différentiable $\mathscr{X} =]a, b[$ par rapport à la densité $\mathrm{d}s$ définie dans la carte $(\mathscr{X}, \mathrm{id})$ par la formule $t \mapsto |\,\mathbf{r}'(t)\,|$. Dans le cas particulier où la courbe γ (ou plus exactement sa restriction à $]a,b[$) est simple et régulière, i.e. lorsque son support $\mathscr{L}$ est une sous-variété plongée à une dimension et le couple $(\mathscr{L}, \gamma^{-1})$, une carte sur $\mathscr{L}$, l'intégrale (13) peut être traitée aussi comme une intégrale sur $\mathscr{L}$ de la fonction f (plus exactement, de sa restriction à $\mathscr{L}$) par rapport à la densité volumique $\mathrm{d}s$ (appel e alors *élément de longueur*), définie dans la carte $(\mathscr{L}, \gamma^{-1})$ par la fonction $t \mapsto |\,\mathbf{r}'(t)\,|$.

LEÇON 25

Variétés orientables. — Intégration des formes. — Lemme de Poincaré pour formes à support borné. — Le groupe $H^n_{\text{s.b.}}\mathscr{X}$. — Cas d'une variété orientable.

Pour appliquer le calcul intégral aux groupes de cohomologie de de Rham, il faut apprendre à intégrer les formes et non pas les densités. A cet effet il faut étendre la notion d'orientation d'un espace vectoriel (ou affine) développée au premier ouvrage (voir *Leçons de géométrie. Géométrie analytique*) aux variétés différentiables arbitraires.

Définition 1. Deux cartes $(U, h) = (U, x^1, \ldots, x^n)$ et $(U', h') = (U', x^{1'}, \ldots, x^{n'})$ d'une variété différentiable $\mathscr{X}$ à n dimensions $(n > 0)$ sont dites *positivement compatibles* si soit $U \cap U' = \varnothing$, soit $U \cap U' \neq \varnothing$ et

$$\det \frac{\partial h'}{\partial h} > 0 \quad \text{sur } U \cap U',$$

i.e. si en chaque point $p \in U \cap U'$ les bases

$$\left(\frac{\partial}{\partial x^1}\right)_p, \ \ldots, \ \left(\frac{\partial}{\partial x^n}\right)_p \quad \text{et} \quad \left(\frac{\partial}{\partial x^{1'}}\right)_p, \ \ldots, \ \left(\frac{\partial}{\partial x^{n'}}\right)_p$$

de l'espace tangent $\mathbf{T}_p\mathscr{X}$ sont de même sens. Un atlas composé de cartes positivement compatibles s'appelle *orienteur*. Une variété est dite *orientable* si elle est munie d'un atlas orienteur au moins.

Il est clair qu'*une variété est orientable si et seulement s'il en est de même de toutes ses composantes.*

Il est aisé de voir (cf. corollaire 1 de la proposition 1 de la leçon 6) que pour tout atlas orienteur A d'une variété orientable $\mathscr{X}$ l'ensemble $\mathsf{A}^+_{\max}$ de toutes les cartes positivement compatibles avec chaque carte de l'atlas A est un atlas orienteur qui contient A et qui de plus est maximal (i.e. tel que si un atlas orienteur A^* contient l'atlas A, alors $\mathsf{A}^* \subset \mathsf{A}^+_{\max}$).

Exercice 1. Montrer que

a) les sphères $\mathbf{S}^n$, $n \geqslant 0$,

b) les espaces projectifs $\mathbb{R}P^{2n+1}$ de dimension impaire,

c) les décomplexifiés des variétés $\mathbb{C}$-analytiques (cf. leçon 11),

d) les groupes de Lie

sont orientables.

[On démontre (non sans peine !) que les espaces projectifs de

dimension paire, et en particulier le plan projectif $\mathbb{R}P^2$, ne sont pas orientables.]

Définition 2. Une variété orientable dans laquelle on a choisi un atlas orienteur maximal est dite *orientée* et l'atlas choisi est son *orientation*. Les cartes appartenant à l'orientation d'une variété orientée sont dites *positivement orientées* (ou tout simplement *positives*).

Par définition, pour tout point $p \in \mathscr{X}$ l'orientation de l'espace tangent $\mathbf{T}_p\mathscr{X}$ définie par la base

$$\left(\frac{\partial}{\partial x^1}\right)_p, \ldots, \left(\frac{\partial}{\partial x^n}\right)_p$$

est la même pour toutes les cartes positivement orientées $(U, x^1, \ldots, x^n)$. On dit de cette orientation qu'elle est *induite* par celle de la variété $\mathscr{X}$.

Donc, de façon imagée, orienter une variété c'est choisir des orientations compatibles sur ses espaces tangents.

Soit $(U, h) = (U, x^1, x^2, \ldots, x^n)$ une carte quelconque d'une variété orientée $\mathscr{X}$. On dira qu'un point $p \in U$ est *positif* si pour une (et donc pour toute) carte positivement orientée (V, k) telle que $p \in V$ on a

$$\det \frac{\partial h}{\partial k} > 0.$$

Il est clair que l'ensemble de tous les points positifs et son complémentaire sont ouverts dans U. Donc, si une carte (U, h) est *connexe* (i.e. l'ensemble U est connexe), ou bien tous les points de U sont positifs et alors la carte (U, h) est positivement orientée, ou bien il n'existe aucun point positif dans U et alors c'est la carte $(U, -x^1, x^2, \ldots, x^n)$ qui est positivement orientée. Donc, *pour toute carte connexe* $(U, x^1, x^2, \ldots, x^n)$ *d'une variété orientée* $\mathscr{X}$, *une (et une seule) des deux cartes* $(U, x^1, x^2, \ldots, x^n)$ *et* $(U, -x^1, x^2, \ldots, x^n)$ *est positivement orientée.*

Exercice 2. En déduire qu'*il n'existe que deux orientations sur une variété orientable connexe de dimension* $n > 0$. [Ces orientations sont dites *contraires*. On désigne généralement par $-\mathscr{X}$ la variété $\mathscr{X}$ munie de l'orientation contraire.]

Il est clair que les diverses composantes d'une variété orientable peuvent être orientées indépendamment l'une de l'autre. Donc, sur une variété $\mathscr{X}$ à N composantes, il existe exactement 2^N orientations différentes.

* * *

La proposition suivante nous donne une idée de l'importance des variétés orientées en théorie de l'intégration:

Proposition 1. *Sur une variété orientée* $\mathscr{X}$ *à* n *dimensions* $(n > 0)$ *il existe une correspondance bijective naturelle entre les densités différentiables et les formes différentielles différentiables de degré* n.

D é m o n s t r a t i o n. Soit ρ une densité différentiable sur $\mathscr{X}$. Pour toute carte positive $(U, h) = (U, x^1, \ldots, x^n)$ définissons une forme ω_U de degré n sur U à l'aide de la formule

$$\omega_U = \rho^U \, dx^1 \wedge \ldots \wedge dx^n.$$

Etant donné que pour deux cartes (U, h) et (U', h') positives (donc positivement compatibles) on a

$$\rho^{U'} = \rho^U \det \frac{\partial h'}{\partial h} \quad \text{sur } U \cap U',$$

il vient $\omega_U = \omega_{U'}$ sur $U \cap U'$, et par suite, la formule

$$\omega = \omega_U \quad \text{sur } U$$

définit de façon unique une forme différentielle ω de degré n sur $\mathscr{X}$.

Réciproquement, soient ω une forme différentielle arbitraire de degré n sur $\mathscr{X}$, $(U, h) = (U, x^1, x^2, \ldots, x^n)$ une carte dans $\mathscr{X}$, et

$$\omega = \omega^U \, dx^1 \wedge \ldots \wedge dx^n \quad \text{sur } U.$$

Définissons une fonction ρ^U sur U en admettant que sur chaque composante U' de l'ensemble U

$$(1) \quad \rho^U = \begin{cases} w^U \text{ si la carte } (U', x^1, x^2, \ldots, x^n) \text{ est positivement orientée,} \\ -w^U \text{ dans le cas contraire (i.e. si c'est la carte } (U', -x^1, x^2, \ldots, x^n) \text{ qui est positivement orientée).} \end{cases}$$

Une vérification évidente montre que les fonctions ρ^U forment une densité et que les correspondances $\rho \mapsto \omega$ et $\omega \mapsto \rho$ sont réciproques l'une de l'autre. □

Les formes de degré n sur une variété à n dimensions s'appellent aussi *formes de degré maximal*.

Pour toute forme ω de degré n sur une variété $\mathscr{X}$ orientée (paracompacte et séparée) à n dimensions et tout ensemble cubable D posons

$$\int_D \omega = \int_D \rho,$$

où ρ est la densité (1) associée à la forme ω. De façon analogue, si une forme ω est *à support borné* (i.e. nulle à l'extérieur d'un ensemble compact, ou, ce qui est équivalent, si la densité ρ qui lui est associée est à support borné), alors par définition

$$\int \omega = \int \rho.$$

Comme pour les densités, les deux formes d'intégrales se ramènent directement l'une à l'autre: pour toute forme ω à support borné

$$\int \omega = \int_D \omega,$$

où D est un ensemble cubable à l'extérieur duquel la forme ω est nulle, et réciproquement

$$\int_D \omega = \int \chi_D \omega,$$

pour toute forme ω et tout ensemble cubable D.

Pour les intégrales de formes la formule (11) de la leçon 24 s'écrit

$$\int \omega = \sum_\alpha \int_{h_\alpha(U_\alpha)} (\eta_\alpha w^U \circ h_\alpha^{-1})(\mathbf{x})\, d\mathbf{x}, \tag{2}$$

où $\{(U_\alpha, h_\alpha)\}$ est un recouvrement localement fini numérotable composé de cartes positivement orientées, et $\{\eta_\alpha\}$ la partition de l'unité subordonnée à ce recouvrement.

Signalons que les intégrales $\int_D \omega$ et $\int \omega$ dépendent de l'orientation de la variété $\mathscr{X}$ (et par exemple changent de signe lorsqu'on change l'orientation d'une variété connexe).

Exercice 3. Soit donnée sur une variété $\mathscr{X}$ une famille de formes $\{\omega_t\}$ de degré n dépendant différentiablement (resp. continûment) d'un paramètre t. Montrer que l'intégrale

$$\int \omega_t$$

sera une fonction différentiable (resp. continue) de t.

Remarque 1. Dans tout ce qui précède nous avons admis que $n > 0$. Dans le cas dégénéré $n = 0$ la variété $\mathscr{X}$ est un ensemble de points isolés et l'orientation doit être spécialement définie. On dira qu'une variété 0-dimensionnelle $\mathscr{X}$ est *orientée* si chacun de ses points p est affecté du signe $\varepsilon(p) = \pm 1$. Sont des formes de degré maximal sur une variété 0-dimensionnelle $\mathscr{X}$ toutes les fonctions f définies sur $\mathscr{X}$, et des formes à support borné, les fonctions non nulles uniquement en un nombre fini de points $p \in \mathscr{X}$. L'intégrale d'une telle fonction sur une variété orientée 0-dimensionnelle $\mathscr{X}$ est définie par la formule

$$\int f = \sum_{p \in \mathscr{X}} \varepsilon(p) f(p), \tag{3}$$

qui a un sens puisqu'on a admis que f est à support borné.

* * *

Utilisons en premier lieu les intégrales de formes pour calculer le groupe $H^n\mathscr{X}$ d'une variété $\mathscr{X}$ différentiable compacte séparée orientable à n dimensions. En fait, on calculera —en admettant que $\mathscr{X}$ n'est que paracompacte—non pas le groupe $H^n\mathscr{X}$ mais le groupe $H^n_{\text{s.b}}\,\mathscr{X}$ qui sera confondu avec $H^n\mathscr{X}$ lorsque $\mathscr{X}$ est compacte. Pour cela il nous faut préalablement étendre le lemme de Poincaré aux formes à support borné définies sur $\mathbb{R}^n$.

Toute forme différentielle ω de degré m sur $\mathbb{R}^n$ s'exprime à l'aide de la formule

$$\omega = \sum_{1\leqslant i_1<\ldots<i_m\leqslant n} \ldots \sum w_{i_1\ldots i_m}\, dx^{i_1}\wedge \ldots \wedge dx^{i_m}, \tag{4}$$

où $w_{i_1\ldots i_m}$ sont des fonctions différentiables sur $\mathbb{R}^n$. Considérons les *familles* $\{\omega_p\}$ de formes (4) dépendant d'un point p d'une variété différentiable $\mathscr{X}$. Chaque forme ω_p s'exprime à l'aide de la formule (4) avec des coefficients dépendant de p, i.e. des coefficients qui sont des fonctions sur le produit $\mathbb{R}^n \times \mathscr{X}$. On admettra toujours que *toutes ces fonctions sont différentiables.* [De telles familles $\{\omega_p\}$ s'identifient de façon naturelle aux formes sur $\mathbb{R}^n \times \mathscr{X}$ dépendant des seules différentielles des coordonnées $x^1, \ldots, x^n$ (cf. leçon 20), mais nous n'aurons pas besoin ici de cette identification.]

On s'intéressera essentiellement aux familles de formes à support borné de degré maximal n. Une telle forme ω_p s'écrit

$$\omega_p = w_p\, dx^1 \wedge \ldots \wedge dx^n, \tag{5}$$

où w_p est une fonction différentiable de $\mathbf{x}$ (et de p) nulle à l'extérieur d'un cube

$$\mathring{I}^n_r = \{\mathbf{x}\in\mathbb{R}^n;\ |x^1|<r,\ \ldots,\ |x^n|<r\}$$

(dont le côté $2r$ dépend généralement de p). Ceci étant,

$$\int_{\mathbb{R}^n} \omega_p = \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} w_p(\mathbf{x})\, d\mathbf{x}. \tag{6}$$

Proposition 2. *Si pour une forme à support borné* (5) *l'intégrale* (6) *est identiquement nulle* (*en* p), *il existe sur* $\mathbb{R}^n$ *des formes à support borné* θ_p *de degré* $n-1$ *dépendant différentiablement du point* $p\in\mathscr{X}$, *telles que*

$$d\theta_p = \omega_p \text{ pour tout point } p\in\mathscr{X}.$$

Ceci étant, on peut admettre subsidiairement que si $\omega_p = 0$ *à l'extérieur de* $\mathring{I}^n_r$, *il en est de même de* θ_p (*en particulier, si* $\omega_p = 0$, *il en est de même de* θ_p).

D é m o n s t r a t i o n. Effectuons une récurrence sur la dimension n. Pour $n = 1$ la forme ω_p s'exprime par la formule

$$\omega_p = w_p \, dx,$$

où $w_p = w_p(x)$ est une fonction différentiable de $x \in \mathbb{R}$ (dépendant différentiablement de p), nulle à l'extérieur d'un intervalle $]-r, r[$. Soit

$$\theta_p(x) = \int_{-r}^{x} w_p(x) \, dx.$$

La fonction θ_p dépend différentiablement de p, et $d\theta_p = w_p$. Comme $\int \omega_p = \theta_p(r)$, on a par hypothèse $\theta_p(r) = 0$, et par suite, $\theta_p(x) = 0$ pour $x \geqslant r$. Puisque de toute évidence $\theta_p(x) = 0$ pour $x \leqslant -r$, on en déduit que $\theta_p(x) = 0$ pour $|x| \geqslant r$. Ceci prouve la proposition 1 pour $n = 1$.

Supposons maintenant que la proposition 1 a été prouvée pour des familles de formes de degré $n - 1$ sur l'espace $\mathbb{R}^{n-1}$ et considérons une famille de formes (5) sur $\mathbb{R}^n$. Soit $p \in \mathscr{X}$ et soit r un nombre (dépendant de p) tel que $\omega_p = 0$ à l'extérieur de $\mathring{I}_r^n$.

Représentons l'espace $\mathbb{R}^n$ par le produit $\mathbb{R}^{n-1} \times \mathbb{R}$. En identifiant par voie de conséquence les points de $\mathbb{R}^n$ aux couples $(\mathbf{x}, t)$, où $\mathbf{x} \in \mathbb{R}^{n-1}$ et $t \in \mathbb{R}$, on peut mettre chaque ω_p sous la forme

$$\omega_p = \omega_{t,p} \wedge dt,$$

où

$$\omega_{t,p} = w_p(\mathbf{x}, t) \, dx^1 \wedge \ldots \wedge dx^{n-1}, \quad \mathbf{x} = (x^1, \ldots, x^{n-1}),$$

est une forme sur $\mathbb{R}^{n-1}$ dépendant différentiablement d'un point (t, p) de la variété $\mathbb{R} \times \mathscr{X}$ et nulle à l'extérieur du cube $\mathring{I}_r^{n-1}$ et pour $|t| \geqslant r$.

La fonction α du lemme 1 de la leçon 1 nous permet de construire immédiatement une fonction α différentiable positive sur $\mathbb{R}^{n-1}$, telle que $\alpha = 0$ à l'extérieur du cube $\mathring{I}_r^{n-1}$, et $\alpha(0) \neq 0$. L'intégrale d'une telle fonction

$$\int_{\mathbb{R}^{n-1}} \alpha(\mathbf{x}) \, d\mathbf{x} \tag{7}$$

est strictement positive, et en divisant la fonction α par cette intégrale, on obtient une fonction α_0 sur $\mathbb{R}^{n-1}$ pour laquelle l'intégrale (7) est égale à 1. Donc, pour la fonction $g_{t,p}$ définie sur $\mathbb{R}^{n-1}$ par la formule

$$g_{t,p}(\mathbf{x}) = w_p(\mathbf{x}, t) - \alpha_0(\mathbf{x}) \int_{\mathbb{R}^{n-1}} w_p(\mathbf{x}, t) \, d\mathbf{x},$$

on a

$$\int_{\mathbb{R}^{n-1}} g_{t,p}(\mathbf{x})\,\mathrm{d}\mathbf{x} = 0.$$

La forme sur $\mathbb{R}^{n-1}$

$$\varpi_{t,p} = g_{t,p}\,\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^{n-1}$$

est nulle à l'extérieur de $\mathring{I}_r^{n-1}$ (et pour $|t| \geqslant r$), dépend différentiablement de (t, p) et l'intégrale de cette forme sur $\mathbb{R}^{n-1}$ est identiquement nulle. Donc, par hypothèse de la récurrence, il existe sur $\mathbb{R}^{n-1}$ une forme $\theta_{t,p}$ dépendant différentiablement de $(t, p) \in \mathbb{R} \times \mathscr{X}$ et nulle à l'extérieur du cube $\mathring{I}_r^{n-1}$ (et pour $|t| \geqslant r$), telle que

$$\mathrm{d}\theta_{t,p} = \varpi_{t,p}.$$

Les formes $\theta_{t,p}$ et $\varpi_{t,p}$ peuvent être traitées comme des formes sur $\mathbb{R}^n$ indépendantes de $\mathrm{d}t$. En tant que telles on les désignera par ϑ_p et ϖ_p respectivement. La forme ϖ_p est liée à la forme ω_p par l'égalité

$$\omega_p = \varpi_p \wedge \mathrm{d}t + f_p\,\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^{n-1} \wedge \mathrm{d}t,$$

où f_p est une fonction sur $\mathbb{R}^n$ définie par la formule

$$f_p(\mathbf{x}, t) = \alpha_0(\mathbf{x}) \int_{\mathbb{R}^{n-1}} \omega_p(\mathbf{x}, t)\,\mathrm{d}\mathbf{x} = w_p(\mathbf{x}, t) - g_{t,p}(\mathbf{x}),$$

et la différentielle $\mathrm{d}\vartheta_p$ de la forme ϑ_p est la somme de la forme ϖ_p et d'une forme $\beta \wedge \mathrm{d}t$. Donc,

$$\mathrm{d}\vartheta_p \wedge \mathrm{d}t = \varpi_p \wedge \mathrm{d}t$$

et, par suite,

$$\mathrm{d}(\vartheta_p \wedge \mathrm{d}t) = \varpi_p \wedge \mathrm{d}t.$$

Définissons la forme θ_p sur $\mathbb{R}^n$ par la formule

$$\theta_p = \vartheta_p \wedge \mathrm{d}t + (-1)^{n-1} F_p\,\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^{n-1},$$

où F_p est une fonction sur $\mathbb{R}^n$ définie par

$$F_p(\mathbf{x}, t) = \int_{-r}^{t} f_p(\mathbf{x}, t)\,\mathrm{d}t.$$

Alors

$$\begin{aligned}\mathrm{d}\theta_p &= \varpi_p \wedge \mathrm{d}t + (-1)^{n-1} f_p(\mathbf{x}, t)\,\mathrm{d}t \wedge \mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^{n-1} = \\ &= \varpi_p \wedge \mathrm{d}t + f_p(\mathbf{x}, t)\,\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^{n-1} \wedge \mathrm{d}t = \omega_p.\end{aligned}$$

La forme θ_p est visiblement nulle à l'extérieur de $\mathring{I}_r^{n-1} \times \mathbb{R}$. D'autre part, comme

$$\int_{-r}^{r} \int_{\mathbb{R}^{n-1}} w_p(\mathbf{x}, t)\, d\mathbf{x}\, dt = \int \omega_p = 0,$$

on a $F_p(\mathbf{x}, t) = 0$ pour $t \geqslant r$ (et bien sûr pour $t \leqslant -r$). Par ailleurs, vu que $\omega_{t,p} = 0$ pour $|t| \geqslant r$, on a $\theta_{t,p} = 0$ pour $|t| \geqslant r$, i.e. $\vartheta_p = 0$ pour $|t| \geqslant r$. Donc, $\theta_p = 0$ pour $|t| \geqslant r$, i.e. $\theta_p = 0$ à l'extérieur du cube $\mathring{I}_r^n$. □

Corollaire 1. *Si pour une forme à support borné ω de degré n sur $\mathbb{R}^n$ on a*

$$\int_{\mathbb{R}^n} \omega = 0, \tag{8}$$

il existe alors sur $\mathbb{R}^n$ une forme θ à support borné de degré $n-1$, telle que

$$\omega = d\theta. \tag{9}$$

Ceci étant, si $\omega = 0$ à l'extérieur du cube $\mathring{I}_r^n$, on peut admettre subsidiairement que $\theta = 0$ à l'extérieur de $\mathring{I}_r^n$. □

Remarque 2. A la leçon suivante on montrera que la condition (8) est non seulement suffisante, mais nécessaire pour que (9) ait lieu.

* * *

Supposons maintenant que $\mathscr{X}$ est une variété (séparée et paracompacte) à n dimensions et soit (U, h) une carte dans $\mathscr{X}$. On dira qu'une forme ω à support borné de degré n sur $\mathscr{X}$ est *concentrée* (sous-entendu sur U) si $\omega = 0$ à l'extérieur de U. Pour une telle forme ω, est définie sur un ensemble ouvert $h(U) \subset \mathbb{R}^n$ la forme $(h^{-1})^* \omega$. Cette forme est nulle à l'extérieur d'un ensemble compact $C \subset h(U)$, d'où il résulte que si on la prolonge à 0 à l'extérieur de $h(U)$, on obtient une forme différentiable (et à support borné) sur $\mathbb{R}^n$ tout entier. Ce prolongement sera aussi désigné par $(h^{-1})^* \omega$.

Soit

$$I\omega = \left| \int_{\mathbb{R}^n} (h^{-1})^* \omega \right|.$$

Le nombre $I\omega$ dépend uniquement de la forme ω, car pour tout autre choix du difféomorphisme h l'intégrale peut tout au plus changer de signe.

[Si la variété $\mathscr{X}$ est orientée, hypothèse que nous préférons ne pas faire pour l'instant, le nombre $I\omega$ n'est autre que la valeur absolue de l'intégrale $\int \omega$.]

On dira qu'une forme ω concentrée à support borné est *essentielle* si $I\omega \neq 0$, *inessentielle* sinon. Une forme essentielle telle que $I\omega = 1$ sera appelée *forme normée*.

Une forme ω à support borné sur $\mathscr{X}$ sera dite *finiment cohomologue à zéro* si sur $\mathscr{X}$ il existe une forme θ à support borné telle que

$$d\theta = \omega.$$

Lemme 1. *Toute forme ω inessentielle concentrée à support borné est finiment cohomologue à zéro.*

Démonstration. Il est clair que sans perdre en généralité on peut admettre que l'ensemble $h(U) \subset \mathbb{R}^n$ est le cube $\mathring{I}^n$, donc que la forme $(h^{-1})^*\omega$ (traitée comme une forme sur l'espace $\mathbb{R}^n$ tout entier) est nulle à l'extérieur du cube $\mathring{I}^n$. Ceci étant, par hypothèse,

$$\int_{\mathbb{R}^n} (h^{-1})^*\omega = 0.$$

Donc, en vertu du corollaire 1 de la proposition 2, il existe sur $\mathbb{R}^n$ une forme θ à support borné, nulle à l'extérieur de $\mathring{I}^n$, telle que

$$(h^{-1})^*\omega = d\theta.$$

Considérons la forme $h^*\theta$. Cette forme est définie sur U et satisfait sur U à la relation

$$d(h^*\theta) = \omega.$$

Par ailleurs, la forme $h^*\theta$ est nulle à l'extérieur d'un sous-ensemble compact (image par h^{-1} d'un sous-ensemble compact du cube $\mathring{I}^n = h(U)$ à l'extérieur duquel la forme θ est nulle) du voisinage de coordonnées U. Donc, en prolongeant cette forme à 0 à l'extérieur de U, on obtient une forme différentielle θ_1 à support borné, différentiable sur $\mathscr{X}$ qui satisfait à la relation

$$d\theta_1 = \omega$$

sur U et à l'extérieur de U, i.e. sur $\mathscr{X}$ tout entière. □

Lemme 2. *Pour tout voisinage de coordonnées $U \subset \mathscr{X}$ il existe sur $\mathscr{X}$ une forme ω normée à support borné, concentrée sur U.*

Démonstration. En vertu de la proposition 2 de la leçon 14, il existe sur $\mathscr{X}$ une fonction φ différentiable positive, nulle à l'extérieur de U et égale à 1 sur un ensemble ouvert $W \subset U$ à adhérence compacte $\overline{W} \subset U$. Définissons sur $\mathscr{X}$ la forme $p \mapsto \omega_p$ en

posant

$$\omega_p = \begin{cases} \varphi(p)\, dx_p^1 \wedge \ldots \wedge dx_p^n & \text{si } p \in U, \\ 0 & \text{sinon,} \end{cases}$$

où $x^1, \ldots, x^n$ sont des coordonnées locales sur U. Il est clair que la forme ω est différentiable à support borné et concentrée sur U. De plus,

$$(10) \qquad \int_{h(U)} (h^{-1})^* \omega = \int_{h(U)} (\varphi \circ h^{-1})(\mathbf{x})\, d\mathbf{x} >$$

$$> \int_{h(\overline{W})} (\varphi \circ h^{-1})(\mathbf{x})\, d\mathbf{x} = \int_{h(\overline{W})} d\mathbf{x} > 0,$$

où $h : U \to h(U)$ est un homéomorphisme de coordonnées. Donc, la forme ω est essentielle. En la divisant par l'intégrale $I\omega$, on obtient une forme normée concentrée sur U. □

Remarque 3. Si la variété $\mathscr{X}$ est orientée et la carte $(U, x^1, \ldots, x^n)$ est positive, la forme ω construite est telle que

$$\int \omega > 0.$$

On dira que deux voisinages de coordonnées U et V dans une variété $\mathscr{X}$ sont *enchaînés* si dans $\mathscr{X}$ il existe des voisinages de coordonnées

$$U_0, U_1, \ldots, U_m,$$

tels que $U_0 = U$, $U_m = V$ et $U_{i-1} \cap U_i \neq \varnothing$ pour tout $i = 1, \ldots, m$ (cf. leçon 20). Il est aisé de voir (cf. lemme 1 de la leçon 20) que *sur une variété connexe $\mathscr{X}$ les voisinages de coordonnées sont deux à deux enchaînés.*

On dira que des formes ω_0 et ω_1 à support borné sur une variété $\mathscr{X}$ sont *finiment cohomologues* s'il existe une forme θ à support borné, telle que

$$\omega_1 - \omega_0 = d\theta,$$

i.e. si la forme $\omega_1 - \omega_0$ est finiment cohomologue à 0.

Proposition 3. *Soient ω_0 et ω_1 des formes à support borné concentrées et normées de degré n sur une variété connexe $\mathscr{X}$ à n dimensions. La forme ω_1 est alors finiment cohomologue soit à ω_0, soit à $-\omega_0$.*

Démonstration. Supposons que la forme ω_0 est concentrée sur un voisinage de coordonnées U_0, et la forme ω_1, sur un voisinage de coordonnées U_1, et soient $h_0 : U_0 \to \mathbb{R}^n$ et $h_1 : U_1 \to \mathbb{R}^n$ des applications de coordonnées telles que

$$(11) \qquad \int_{\mathbb{R}^n} (h_0^{-1})^* \omega_0 = 1 \quad \text{et} \quad \int_{\mathbb{R}^n} (h_1^{-1})^* \omega_1 = 1.$$

Cas 1. Les cartes (U_0, h_0) et (U_1, h_1) sont confondues. La forme $\omega_1 - \omega_0$ à support borné est alors concentrée sur U_0 et est inessentielle. Donc, $\omega_1 - \omega_0$ est finiment cohomologue à 0 en vertu du lemme 1.

Cas 2. Les voisinages de coordonnées U_0 et U_1 sont confondus. Les intégrales (11) pouvant tout au plus changer de signe lorsqu'on modifie les applications de coordonnées, on a

$$\int_{\mathbb{R}^n} (h_1^{-1})^* \omega_0 = \varepsilon, \quad \text{où} \quad \varepsilon = \pm 1.$$

Donc. la forme ω_1 est finiment cohomologue à la forme $\varepsilon\omega_0$ d'après ce qui a été déjà démontré.

Cas 3. Les voisinages de coordonnées U_0 et U_1 se rencontrent:

$$U_0 \cap U_1 \neq \varnothing.$$

D'après le lemme 2, il existe une forme normée ω concentrée sur le voisinage de coordonnées $U_0 \cap U_1$ (donc sur les voisinages de coordonnées U_0 et U_1). Les formes normées ω_0 et ω étant concentrées sur U_0, d'après ce qui vient d'être prouvé la forme ω_0 est finiment cohomologue soit à la forme ω, soit à la forme $-\omega$, et comme les formes normées ω et ω_1 sont concentrées sur U_1, la forme ω est finiment cohomologue soit à la forme ω_1, soit à la forme $-\omega_1$. Donc, la forme ω_1 est finiment cohomologue soit à la forme ω_0, soit à la forme $-\omega_0$.

Cas 4. Les voisinages de coordonnées U_0 et U_1 sont enchaînés. Par une récurrence évidente sur la longueur de la chaînette de voisinages de coordonnées reliant les voisinages U_0 et U_1 on ramène ce cas au cas 3 déjà étudié.

Ceci prouve la proposition 2, puisque dans une variété connexe les voisinages de coordonnées sont deux à deux enchaînés. □

Corollaire 1. *Soit ω_0 une forme concentrée essentielle de degré n sur une variété connexe $\mathscr{X}$ à n dimensions. Pour toute forme ω à support borné de degré n sur $\mathscr{X}$ il existe alors un nombre c tel que la forme ω soit finiment cohomologue à la forme $c\omega_0$.*

Démonstration. Sans perdre en généralité on peut admettre que la forme ω_0 est normée.

La variété $\mathscr{X}$ étant par hypothèse séparée et paracompacte, il existe sur elle un recouvrement numérotable $\{U_\alpha\}$ composé de voisinages de coordonnées. Soit $\{\eta_\alpha\}$ une partition de l'unité subordonnée au recouvrement $\{U_\alpha\}$. Pour tout α la forme $\eta_\alpha\omega$ est alors concentrée sur U_α, donc est défini le nombre

$$c_\alpha = I(\eta_\alpha\omega).$$

Si $c_\alpha \neq 0$, la forme $c_\alpha^{-1}\eta_\alpha\omega$ est normée et concentrée tout comme la forme ω_0. Donc, en vertu de la proposition 3, cette forme est finiment cohomologue à la forme $\varepsilon_\alpha\omega_0$, où $\varepsilon_\alpha = \pm 1$. Par consé-

quent, la forme $\eta_\alpha\omega$ est finiment cohomologue à la forme $c_\alpha\varepsilon_\alpha\omega_0$.

Il est clair que ceci est valable pour $c_\alpha = 0$, car la forme concentrée $\eta_\alpha\omega$ est alors inessentielle et, par suite, finiment cohomologue à 0 en vertu du lemme 2.

Donc, dans le développement

$$\omega = \sum_\alpha \eta_\alpha\omega$$

(qui ne contient qu'un nombre fini de termes non nuls, puisque la forme ω est à support borné), chaque terme $\eta_\alpha\omega$ est finiment cohomologue à la forme $c_\alpha\varepsilon_\alpha\omega_0$. Donc, la forme ω est finiment cohomologue à la forme $c\omega_0$, où

$$c = \sum_\alpha c_\alpha\varepsilon_\alpha. \quad \square$$

Pour tout $m \geqslant 0$ posons

$$H^m_{\text{s.b.}}\mathscr{X} = Z^m_{\text{s.b.}}\mathscr{X}/B^m_{\text{s.b.}}\mathscr{X},$$

où $Z^m_{\text{s.b.}}\mathscr{X}$ est l'espace de toutes les formes fermées à support borné de degré m sur $\mathscr{X}$, et $B^m_{\text{s.b.}}\mathscr{X}$ le sous-espace de $Z^m_{\text{s.b.}}\mathscr{X}$ composé des différentielles des formes à support fini de degré $m - 1$. (On remarquera que d'une façon générale il existe des formes à support borné exactes n'appartenant pas à $B^m_{\text{s.b.}}\mathscr{X}$.)

Les éléments de l'espace quotient $H^m_{\text{s.b.}}\mathscr{X}$ s'appellent *classes de cohomologie à support borné* de la variété différentiable $\mathscr{X}$.

Théorème 1. *Pour toute variété $\mathscr{X}$ connexe séparée et paracompacte à n dimensions on a l'inégalité*

$$\dim H^n_{\text{s.b.}}\mathscr{X} \leqslant 1,$$

i.e. soit $H^n_{\text{s.b.}}\mathscr{X} = 0$, soit $H^n_{\text{s.b.}}\mathscr{X} \approx \mathbb{R}$.

D é m o n s t r a t i o n. D'après le corollaire 1 de la proposition 3 l'espace vectoriel $H^n_{\text{s.b.}}\mathscr{X}$ est engendré par la classe de cohomologie à support borné $[\omega_0]$ d'une forme ω_0 essentielle concentrée. Donc, si $[\omega_0] = 0$, alors $H^n_{\text{s.b.}}\mathscr{X} = 0$, et si $[\omega_0] \neq 0$, alors $H^n_{\text{s.b.}}\mathscr{X} \approx \mathbb{R}$. $\square$

* * *

Lorsque la variété $\mathscr{X}$ est orientable, on peut affiner le résultat.

En munissant $\mathscr{X}$ d'une orientation (on rappelle que $\mathscr{X}$ est supposée connexe), on peut pour chaque forme $\omega \in \Omega^n\mathscr{X}$ à support borné construire l'intégrale

$$\int \omega. \tag{12}$$

A la leçon 27 on montrera (cf. corollaire 1 du théorème 1, leçon 27) que *si $\omega \in B^n_{\text{s.b.}}\mathscr{X}$, l'intégrale* (12) *est nulle*. Donc, l'intégrale (12) ne dépend que de la classe de cohomologie à support bor-

né $[\omega] \in H^n_{\text{s.b.}}\mathscr{X}$ de la forme ω (qui, remarquons-le, est automatiquement fermée), i.e. la correspondance

$$[\omega] \mapsto \int \omega \tag{13}$$

définit de façon unique un homomorphisme $H^n_{\text{s.b.}}\mathscr{X} \to \mathbb{R}$.

Théorème 2. *Pour toute variété $\mathscr{X}$ connexe séparée paracompacte orientable à n dimensions le groupe $H^n_{\text{s.b.}}\mathscr{X}$ est isomorphe à $\mathbb{R}$. L'isomorphisme est défini par la correspondance* (13). (*Il dépend de l'orientation choisie dans la variété $\mathscr{X}$.*)

Démonstration. De la remarque 3 on sait que sur la variété $\mathscr{X}$ il existe des formes ω à support borné de degré n pour lesquelles l'intégrale (2) est différente de 0, ce qui exprime que l'homomorphisme (13) n'est pas trivial.

Ceci n'est possible, puisque $\dim H^n_{\text{s.b.}}\mathscr{X} \leqslant 1$, que si $\dim H^n_{\text{s.b.}}\mathscr{X} = 1$, et l'homomorphisme (13) est alors un isomorphisme. □

Donc, pour qu'une classe de cohomologie à support borné $[\omega]$ d'une forme ω à support borné (pas forcément concentrée) de degré n engendre un espace vectoriel $H^n_{\text{s.b.}}\mathscr{X}$ (constitue sa base), il est nécessaire et suffisant que

$$\int \omega \neq 0. \tag{14}$$

Dans ce cas, pour toute autre forme ω_1 à support borné de degré n sur $\mathscr{X}$, on a

$$[\omega_1] = c\,[\omega],$$

où

$$c = \frac{\int \omega_1}{\int \omega}.$$

Remarquons que la condition (14) est visiblement remplie si la forme ω est concentrée et essentielle.

Corollaire 1. *Pour toute variété $\mathscr{X}$ connexe séparée compacte orientable à n dimensions le groupe $H^n\mathscr{X}$ est isomorphe à $\mathbb{R}$, i.e.*

$$h^n\mathscr{X} = 1. \ \square$$

Remarque 4. La donnée de l'isomorphisme (13) équivaut à la donnée d'une base dans l'espace vectoriel à une dimension $H^n_{\text{s.b.}}\mathscr{X}$ et donc à la donnée d'une orientation sur cet espace. Cette orientation change avec celle de la variété $\mathscr{X}$. Donc, *l'orientation de la variété connexe $\mathscr{X}$ peut être identifiée à celle de l'espace vectoriel $H^n_{\text{s.b.}}\mathscr{X}$.*

Remarque 5. On démontre (faites-le à titre d'exercice !) que l'égalité $\dim H^n_{\text{s.b.}}\mathscr{X} = 1$ caractérise les variétés orientables, i.e. *une variété connexe $\mathscr{X}$ n'est pas orientable si et seulement si* $H^n_{\text{s.b.}}\mathscr{X} = 0$.

Exercice 4. Calculer le groupe $H^n_{\text{s.b.}}\mathscr{X}$ pour une variété $\mathscr{X}$ non connexe. Comparer le résultat avec la proposition 2 de la leçon 20.

LEÇON 26

Degré d'une application différentiable propre. — Nombre algébrique des images réciproques d'une valeur régulière. — Invariance du degré par une homotopie différentiable. — Démonstration du théorème du tambour. — Invariance du degré par une homotopie quelconque.

Définition 1. On dit qu'une application $f : \mathscr{X} \to \mathscr{Y}$ est *propre* si l'image réciproque $f^{-1}C$ de chaque ensemble compact $C \subset \mathscr{Y}$ est compacte. Il est clair que si une application différentiable $f : \mathscr{X} \to \mathscr{Y}$ est propre, la forme $f^*\omega$ sur $\mathscr{X}$ est à support borné si la forme ω l'est sur $\mathscr{Y}$. Donc, pour tout $m \geqslant 0$, chaque application différentiable propre $f : \mathscr{X} \to \mathscr{Y}$ induit un homomorphisme

$$f^* : H^m_{\text{s.b.}}\mathscr{Y} \to H^m_{\text{s.b.}}\mathscr{X}.$$

Considérons cet homomorphisme dans le cas particulier où les deux variétés $\mathscr{X}$ et $\mathscr{Y}$ sont orientées et de même dimension ($n = m$). On admettra par ailleurs que *la variété $\mathscr{Y}$ est connexe.*

Supposons qu'une forme ω à support borné de degré n sur $\mathscr{Y}$ n'est pas finiment cohomologue à zéro (sa classe de cohomologie à support borné $[\omega]$ forme une base de l'espace vectoriel $H^n_{\text{s.b.}}\mathscr{Y}$). On sait alors que $\int\limits_{\mathscr{Y}} \omega \neq 0$, et pour toute autre forme ω_1 à support borné de degré n sur $\mathscr{Y}$ on aura l'égalité $[\omega_1] = c\,[\omega]$, où

$$c = \frac{\int\limits_{\mathscr{Y}} \omega_1}{\int\limits_{\mathscr{Y}} \omega}.$$

Les formes ω_1 et $c\omega$ étant finiment cohomologues, il en est de même des formes $f^*\omega_1$ et $f^*(c\omega) = cf^*\omega$, et par suite, lorsque $[f^*\omega] \neq 0$ dans $H^n_{\text{s.a.}}\mathscr{X}$, on a

$$c = \frac{\int\limits_{\mathscr{X}} f^*\omega_1}{\int\limits_{\mathscr{X}} f^*\omega}.$$

Pour $c \neq 0$, i.e. pour $\int\limits_{\mathscr{Y}} \omega_1 \neq 0$, il s'ensuit que

$$\frac{\int\limits_{\mathscr{X}} f^*\omega}{\int\limits_{\mathscr{Y}} \omega} = \frac{\int\limits_{\mathscr{X}} f^*\omega_1}{\int\limits_{\mathscr{Y}} \omega_1},$$

i.e. le nombre

$$\deg f = \frac{\int_{\mathscr{X}} f^*\omega}{\int_{\mathscr{Y}} \omega} \tag{1}$$

est indépendant du choix de la forme ω.

Il est clair que ceci vaut également pour $f^*[\omega]=0$.

Remarque 1. La justification algébrique de ce raisonnement réside dans le fait évident que chaque application linéaire $\mathbb{R} \to \mathbb{R}$ est une multiplication par un nombre.

Définition 2. Le nombre (1) s'appelle *degré* de l'application différentiable propre $f : \mathscr{X} \to \mathscr{Y}$.

L'égalité (1) peut être mise sous la forme

$$\int_{\mathscr{X}} f^*\omega = \deg f \cdot \int_{\mathscr{Y}} \omega, \tag{2}$$

et sous cet aspect elle a lieu pour toute forme ω à support borné de degré maximal sur $\mathscr{Y}$.

Signalons qu'une condition nécessaire pour que $\deg f$ soit défini est que l'application f soit propre. Cette condition est toujours remplie lorsque la variété $\mathscr{X}$ est compacte. Donc, *si la variété $\mathscr{X}$ est compacte*, $\deg f$ *est défini pour toute application différentiable* $f : \mathscr{X} \to \mathscr{Y}$ (à condition bien sûr que $\mathscr{Y}$ soit connexe).

Exercice 1. Montrer que si la variété $\mathscr{X}$ est compacte et la variété $\mathscr{Y}$ ne l'est pas, toute application $\mathscr{X} \to \mathscr{Y}$ est de degré 0.

Remarque 2. Pour que le degré soit défini, il faut aussi que les variétés $\mathscr{X}$ et $\mathscr{Y}$ soient non seulement orientables mais orientées. De plus, le degré change de signe quand on modifie l'orientation de l'une d'elles. Mais pour $\mathscr{Y} = \mathscr{X}$ le changement d'orientation de la variété $\mathscr{X}$ laisse invariant le degré (qui change deux fois de signe). Donc, on peut parler du degré d'une application propre $\mathscr{X} \to \mathscr{X}$ (où $\mathscr{X}$ est une variété connexe séparée et paracompacte) sans admettre que $\mathscr{X}$ soit nécessairement orientée (il faut seulement qu'elle soit orientable).

* * *

D'après le théorème de Sard (cf. leçon 15) l'application différentiable $f : \mathscr{X} \to \mathscr{Y}$ possède des valeurs régulières $q \in \mathscr{Y}$ et, en vertu de la proposition 1 de la leçon 13, pour toute valeur régulière $q \in \mathscr{Y}$ l'ensemble $f^{-1}(q)$ est une sous-variété 0-dimensionnelle plongée dans la variété $\mathscr{X}$ (on rappelle que par hypothèse $\dim \mathscr{Y} = \dim \mathscr{X}$), i.e. est composé de points isolés (cf. leçon 7). Ceci étant, l'application f étant propre, cet ensemble est compact et, par suite, fini.

Soient p un point arbitraire de l'ensemble $f^{-1}(q)$, $(U, h) = (U, x^1, \ldots, x^n)$ une carte positive de la variété $\mathscr{X}$, centrée en p, $(V, k) = (V, y^1, \ldots, y^n)$ une carte positive de la variété $\mathscr{Y}$, centrée en q, et

$$y^j = f^j(x^1, \ldots, x^n), \quad j = 1, \ldots, n, \tag{3}$$

les fonctions exprimant la fonction f dans ces cartes. Par hypothèse, le jacobien

$$D_f(p) = \det \left\| \frac{\partial f^j}{\partial x^i} \right\|, \quad i, j = 1, \ldots, n,$$

des fonctions (3) en p est différent de 0. Donc, son signe $\operatorname{sgn} D_f(p) = \pm 1$ est défini. Il est évident que ce signe ne dépend pas du choix des cartes (positives !) (U, h) et (V, k). On l'appellera *signe de l'application* f au point p et on le désignera par le symbole $\operatorname{sgn}_p f$.

Soit

$$\sum_{f(p)=q} \operatorname{sgn}_p f \tag{4}$$

la somme des signes de l'application f en tous les points $p \in f^{-1}(q)$ (« le nombre algébrique des images réciproques du point q par l'application f »). Cette somme est définie car l'ensemble $f^{-1}(q)$ est fini.

Proposition 1. *Le nombre* (4) *est indépendant du choix du point* $q \in \mathscr{Y}$ *et est égal à* $\deg f$.

Démonstration. Soit W un voisinage du point $q \in \mathscr{Y}$ d'adhérence $\overline{W}$ compacte.

L'application f étant continue et l'ensemble de tous ses points critiques fermé, tout point $p \in f^{-1}(q)$ possède un voisinage $U'_p \subset f^{-1}W$ ne contenant aucun point critique de f. Ceci étant, les voisinages U'_p peuvent visiblement être choisis de telle sorte qu'ils soient disjoints pour des points $p \in f^{-1}(q)$ distincts, que tous les ensembles $f(U'_p)$ soient ouverts dans $\mathscr{Y}$ et que toutes les applications $f \mid_{U'_p} : U'_p \to f(U'_p)$ soient des difféomorphismes.

Soit

$$U' = \bigcup_{f(p)=q} U'_p$$

la réunion de tous les voisinages U'_p et soit $C = f^{-1}\overline{W} \setminus U'$. L'ensemble C est compact et tel que $q \notin f(C)$. Son image $f(C)$ est compacte, donc fermée. Il existe donc sur $\mathscr{Y}$ un voisinage de coordonnées V' de q tel que $V' \subset W \setminus f(C)$ et, par suite, tel que $f^{-1}V' \subset U'$.

Soit

$$V = \bigcap_{f(p)=q} f(U'_p) \cap V'.$$

L'ensemble $f^{-1}(q)$ étant fini, l'ensemble V (qui contient q) est ouvert et, par suite, est un voisinage de q. Pour tout point $p \in f^{-1}(q)$ posons

$$U_p = U'_p \cap f^{-1}V.$$

Il est évident que les ensembles U_p possèdent les propriétés suivantes:

a) Chaque ensemble U_p est un voisinage du point p.

b) Les voisinages U_p correspondant à des points p distincts sont disjoints.

c) La réunion U de tous les voisinages U_p, $p \in f^{-1}(q)$, est l'image réciproque $f^{-1}V$ de V.

d) Sur chaque voisinage U_p l'application f est un difféomorphisme de U_p sur V.

Puisque le voisinage V est contenu dans un voisinage de coordonnées V', il est lui-même un voisinage de coordonnées, i.e. est le support d'une carte $(V, k) = (V, y^1, \ldots, y^n)$ de $\mathcal{Y}$ qui sans perte de généralité peut être supposée positive. D'après la propriété d) on déduit de là que pour tout point $p \in f^{-1}(q)$, le couple (U_p, h_p), où $h_p = k \circ (f \mid_{U_p})$, est une carte aussi. Puisque l'application f s'exprime par des formules de la forme $y^i = x^i$, $i = 1, \ldots, n$, dans les cartes (U_p, h_p) et (V, k), la carte (U_p, h_p) est positive si et seulement si $\operatorname{sgn}_p [f] = 1$.

On sait que pour calculer $\deg f$ on peut utiliser une forme ω à support borné de degré n sur $\mathcal{Y}$ non finiment cohomologue à 0. En usant de cette liberté, on peut prendre pour ω une forme essentielle concentrée sur V. Une telle forme existe en vertu du lemme 2 de la leçon 25.

Mais si la forme ω est concentrée sur V, la forme $f^*\omega$ sera nulle à l'extérieur des voisinages U_p et donc

$$\int_{\mathcal{X}} f^*\omega = \sum_{f(p)=q} \int_{U_p} f^*\omega,$$

où au second membre U_p est traité comme une sous-variété ouverte munie d'une orientation induite par celle de $\mathcal{X}$. Donc, si

$$\omega = w\, \mathrm{d}y^1 \wedge \ldots \wedge \mathrm{d}y^n \quad \text{sur } V$$

et, par suite,

$$f^*\omega = (w \circ f)\, \mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^n \quad \text{sur } U_p,$$

où $x^1 = y^1 \circ f, \ldots, x^n = y^n \circ f$, alors

$$\int_{U_p} f^*\omega = \varepsilon_p \int_{h_p(U_p)} (w \circ f \circ h_p^{-1})(\mathbf{x})\, \mathrm{d}\mathbf{x},$$

où $\varepsilon_p = 1$ si la carte (U_p, h_p) est positive et $\varepsilon_p = -1$ sinon. D'autre part, puisque $h_p = k \circ (f \mid_{U_p})$ et donc $h_p(U_p) = k(V)$ et

$w \circ f \circ h_p^{-1} = w \circ k^{-1}$, alors

$$\int_{h_p(U_p)} (w \circ f \circ h_p^{-1})(\mathrm{x})(\mathrm{dx}) = \int_{k(V)} (w \circ k^{-1})(\mathrm{x})\,\mathrm{dx} = \int_{\mathscr{Y}} \omega.$$

Par ailleurs, d'après la remarque faite ci-dessus, l'égalité $\varepsilon_p = +1$ a lieu si et seulement si $\operatorname{sgn}_p f = 1$. Puisque $\varepsilon_p = \pm 1$ et $\operatorname{sgn}_p f = \pm 1$, ceci exprime que

$$\varepsilon_p = \operatorname{sgn}_p f \quad \text{pour tout point } p \in f^{-1}(q).$$

En collationnant ces faits, on trouve immédiatement que

$$\int_{\mathscr{X}} f^*\omega = \Big(\sum_{f(p)=q} \operatorname{sgn}_p f\Big) \int_{\mathscr{Y}} \omega.$$

Par définition (cf. formule (1)) ceci exprime que le nombre (4) est égal à $\deg f$. □

Corollaire 1. *Le degré de toute application différentiable propre $f: \mathscr{X} \to \mathscr{Y}$ est un nombre entier.* □

Résultat surprenant!

Corollaire 2. *Le degré d'une application $f: \mathscr{X} \to \mathscr{Y}$ non surjective est nul.*

D é m o n s t r a t i o n. Il suffit de remarquer que tout point $q \notin f(\mathscr{X})$ est une valeur régulière de l'application f. □

* * *

Définition 3. On dit que des applications $f, g: \mathscr{X} \to \mathscr{Y}$ continues (resp. continues propres) sont *homotopes* (resp. *proprement homotopes*) s'il existe une application continue (resp. continue propre) $F: \mathscr{X} \times [0, 1] \to \mathscr{Y}$ telle que

$$F(p, 0) = f(p), \quad F(p, 1) = g(p)$$

pour tout point $p \in \mathscr{X}$.

L'application F s'appelle *homotopie* (resp. *homotopie propre*) reliant les applications f et g. Il est commode de l'identifier à la famille $\{f_t\}$ des applications continues (resp. continues propres) $f_t: \mathscr{X} \to \mathscr{Y}$, $0 \leqslant t \leqslant 1$, définies par la formule

$$f_t(p) = F(p, t), \quad p \in \mathscr{X}.$$

On dit qu'une homotopie F est *différentiable* si elle est la restriction à $\mathscr{X} \times [0, 1]$ d'une application différentiable $\mathscr{X}_p \times]a, b[\to \mathscr{Y}$, où $]a, b[$ est un intervalle de l'axe $\mathbb{R}$ contenant le segment $[0, 1]$. Des applications différentiables (et propres) $f, g: \mathscr{X} \to \mathscr{Y}$ reliées par une homotopie différentiable (et propre) $F: \mathscr{X} \times [0, 1] \to \mathscr{Y}$ sont dites *différentiablement* (*et proprement*) *homotopes*.

Etant donné que pour toute homotopie différentiable et propre $F: \mathscr{X} \times [0, 1] \to \mathscr{Y}$ les applications $f_t: \mathscr{X} \to \mathscr{Y}$, $0 \leqslant t \leqslant 1$, sont différentiables et propres, pour chacune d'elles est défini

$$\deg f_t = \frac{\int_{\mathscr{X}} f_t^* \omega}{\int_{\mathscr{Y}} \omega}.$$

La forme $f_t^* \omega$ intervenant dans cette formule dépend de toute évidence continûment (et même différentiablement) de t, i.e. dans chaque carte son unique coefficient est une fonction continue de t (et bien sûr des coordonnées locales). Donc (cf. exercice 3 de la leçon 25) le nombre $\deg f_t$ dépend aussi continûment de t. Il est par conséquent constant, puisque entier. En particulier, $\deg f = \deg f_0 = \deg f_1 = \deg g$.

Donc, *si des applications différentiables propres f, $g: \mathscr{X} \to \mathscr{Y}$ sont différentiablement et proprement homotopes, leurs degrés sont égaux*:

$$\deg f = \deg g. \tag{5}$$

Cette proposition est connue sous le nom de théorème d'invariance homotope du degré.

* * *

Nous pouvons maintenant démontrer sans peine le théorème du tambour annoncé à la leçon 9.

Démonstration du théorème 1 de la leçon 9. Supposons qu'il existe une rétraction

$$r: \mathbf{B}^n \to \mathbf{S}^{n-1}, \quad n \geqslant 2,$$

et soit $r(0) = \mathbf{s}_0$. La formule

$$F(\mathbf{x}, t) = r(t\mathbf{x}), \quad \mathbf{x} \in \mathbf{S}^{n-1}, \quad 0 \leqslant t \leqslant 1, \tag{6}$$

définit alors une homotopie

$$F: \mathbf{S}^{n-1} \times [0, 1] \to \mathbf{S}^{n-1}$$

reliant l'application constante

$$\text{const}: \mathbf{S}^{n-1} \to \mathbf{S}^{n-1}, \quad \mathbf{x} \mapsto \mathbf{s}_0,$$

à l'application identique $\text{id}: \mathbf{S}^{n-1} \to \mathbf{S}^{n-1}$, $\mathbf{x} \mapsto \mathbf{x}$. Donc, en vertu de l'invariance homotope du degré, les applications const et id (qui sont visiblement différentiables) doivent avoir le même degré (peu importe que ces applications soient propres, puisque la sphère $\mathbf{S}^{n-1}$ est compacte). Or il est clair que $\deg \text{id} = 1$ et $\deg \text{const} = 0$. Ce qui contredit l'existence de la rétraction r. □

Remarque 3. Cette démonstration ne passe pas pour $n = 1$ (ne serait-ce que parce que la sphère S^n est composée de deux points et n'est pas une variété connexe). Mais l'impossibilité de l'existence d'une rétraction $B^1 \to S^0$ est évidente dans ce cas (le segment connexe B^1 ne peut être appliqué continûment sur la sphère non connexe S^0).

* * *

Le lecteur attentif aura remarqué que cette démonstration du théorème du tambour renferme une lacune et, par conséquent, est mise à proprement parler en défaut. En effet, l'homotopie (6) n'est généralement que continue, alors que dans l'égalité (5) on a admis que l'homotopie reliant les applications f et g était différentiable. Donc, pour asseoir le théorème du tambour, il nous faut encore prouver (au moins pour les applications $S^{n-1} \to S^{n-1}$) que *l'égalité* (5) *est valable dans le cas aussi où l'homotopie reliant les applications f et g n'est que continue.* Pour cela il suffit de toute évidence de prouver la proposition générale suivante:

Proposition 2. *Pour toutes variétés $\mathscr{X}$ et $\mathscr{Y}$ compactes (et séparées) les applications différentiables $f, g : \mathscr{X} \to \mathscr{Y}$ sont homotopes si et seulement si elles sont différentiablement homotopes.*

[On peut établir une assertion analogue pour les applications propres de variétés non compactes par une complication des détails techniques, mais nous glisserons là-dessus, car nous n'avons besoin du théorème du tambour que dans le cas où $\mathscr{X} = \mathscr{Y} = S^{n-1}$. D'autre part, à strictement parler, on prouvera la proposition 2 seulement sous certaines conditions subsidiaires sur les variétés $\mathscr{X}$ et $\mathscr{Y}$, conditions qui seront visiblement satisfaites pour $\mathscr{X} = \mathscr{Y} = S^{n-1}$. (En fait, elles sont remplies pour toutes $\mathscr{X}$ et $\mathscr{Y}$, mais cette circonstance ne sera établie qu'au prochain ouvrage de cette série.)]

Signalons que dans la proposition 2 les variétés $\mathscr{X}$ et $\mathscr{Y}$ ne sont pas supposées avoir la même dimension (et peuvent même être non connexes).

Les variétés $\mathscr{X}$ et $\mathscr{Y}$ seront supposées plongées dans l'espace $\mathbb{R}^n$, où n est un nombre assez grand. (On renonce donc à l'usage d'utiliser la lettre n pour désigner la dimension de la variété $\mathscr{X}$.) D'après le théorème de plongement (proposition 1 de la leçon 14) cette hypothèse ne restreint pas la généralité.

Définition 4. On dit qu'un sous-ensemble $\mathscr{X}$ de l'espace $\mathbb{R}^n$ est un *retract de voisinage* s'il existe un ensemble ouvert $O \supset \mathscr{X}$ s'appliquant par une rétraction sur $\mathscr{X}$. Ceci étant, dans le cas où $\mathscr{X}$ est une sous-variété, on exige accessoirement qu'il existe une rétraction $r : O \to \mathscr{X}$ qui soit une application différentiable.

Par exemple, *la sphère S^{n-1} de l'espace $\mathbb{R}^n$ est un retract de voisinage.* (Pour voisinage O on peut alors prendre l'espace pointé

$\mathbb{R}^n \setminus \{0\}$ et définir la rétraction $r: O \to S^{n-1}$ par la formule $r(\mathbf{x}) = \frac{\mathbf{x}}{|\mathbf{x}|}$, $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$.)

On prouvera la proposition 2 sous l'hypothèse seulement que *les sous-variétés $\mathscr{X}$ et $\mathscr{Y}$ de l'espace $\mathbb{R}^n$ sont des retracts de voisinage.* [On montrera dans le prochain ouvrage que cette condition ne restreint pas la généralité. De plus elle est remplie en tout cas pour $\mathscr{X} = \mathscr{Y} = S^{n-1}$.]

Démonstration de la proposition 2. Il est évident que si les applications différentiables $f, g: \mathscr{X} \to \mathscr{Y}$ sont différentiablement homotopes, elles sont homotopes. Il nous suffit donc de prouver la réciproque. Un procédé naturel consiste à transformer en une homotopie différentiable l'homotopie continue $F: \mathscr{X} \times [0, 1] \to \mathscr{Y}$ reliant les applications f et g.

Soit $r: O_{\mathscr{X}} \to \mathscr{X}$ une rétraction sur $\mathscr{X}$ d'un voisinage $O_{\mathscr{X}} \supset \mathscr{X}$. Définissons une application

$$F': O_{\mathscr{X}} \times \mathbb{R} \to \mathscr{Y}$$

par la formule

$$F'(\mathbf{x}, t) = \begin{cases} f(r(\mathbf{x})) & \text{si } t \leqslant 3/7, \\ F(r(\mathbf{x}), 7t-3) & \text{si } 3/7 \leqslant t \leqslant 4/7, \\ g(r(\mathbf{x})) & \text{si } 4/7 \leqslant t. \end{cases}$$

Il est clair que l'application F' est continue (et même différentiable pour $t < 3/7$ et $t > 4/7$) et sa restriction à $\mathscr{X} \times [0, 1]$ est une homotopie reliant les applications f et g.

Du cours d'analyse on connaît le

Théorème de Weierstrass (d'approximation polynomiale). *Soient O un ensemble ouvert de l'espace $\mathbb{R}^n$, K un sous-ensemble compact de O et $F: O \to \mathbb{R}^m$ une application continue. Pour tout $\varepsilon > 0$ il existe une application polynomiale $P: \mathbb{R}^n \to \mathbb{R}^m$ (définie par des fonctions polynomiales de coordonnées) telle que*

$$|P(\mathbf{x}) - F(\mathbf{x})| < \varepsilon \quad \textit{pour tout point } \mathbf{x} \in K.$$

[Remarquons qu'il suffit de toute évidence de prouver ce théorème pour $m = 1$ seulement, i.e. dans le cas où F est une fonction numérique. C'est le seul cas généralement envisagé en analyse.]

Appliquons le théorème de Weierstrass à un ensemble ouvert $O_{\mathscr{X}} \times \mathbb{R}$ (traité comme un sous-ensemble de l'espace $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$), à l'ensemble compact $\mathscr{X} \times [0, 1]$ et à l'application F' (traitée comme une application dans $\mathbb{R}^n$). En désignant la restriction de l'application polynomiale P à $\mathscr{X} \times [0, 1]$ encore par P, on trouve en vertu de ce théorème que *pour tout $\varepsilon > 0$ il existe une application différentiable*

$$P: \mathscr{X} \times [0, 1] \to \mathbb{R}^n,$$

telle que

$$|P(\mathbf{x}, t) - F'(\mathbf{x}, t)| < \varepsilon \quad \textit{pour tout point } (\mathbf{x}, t) \in \mathscr{X} \times [0, 1].$$

Soit maintenant λ une fonction différentiable $\mathbb{R} \to \mathbb{R}$ telle que

$$\lambda(t) = \begin{cases} 0 & \text{si } t \leqslant 1/7 \text{ ou } t \geqslant 6/7, \\ 1 & \text{si } 2/7 \leqslant t \leqslant 5/7 \end{cases}$$

et $0 \leqslant \lambda(t) \leqslant 1$ pour tout $t \in \mathbb{R}$. [Une telle fonction peut par exemple être définie par la formule

$$\lambda(t) = \frac{\alpha\left(\frac{5}{14} - \left|\frac{1}{2} + t\right|\right)}{\alpha\left(\frac{5}{14} - \left|\frac{1}{2} + t\right|\right) + \alpha\left(\left|t + \frac{1}{2}\right| - \frac{3}{14}\right)}, \quad -\infty < t < +\infty,$$

où α est la fonction du lemme 1 de la leçon 1 ; cf. remarque 3 de la leçon 1.] Posons

$$G(\mathbf{x}, t) = F'(\mathbf{x}, t) + \lambda(t)(P(\mathbf{x}, t) - F'(\mathbf{x}, t))$$

pour tout point $(\mathbf{x}, t) \in \mathscr{X} \times [0, 1]$.

Puisque $\lambda(t) = 0$ pour $t = 0$ et $t = 1$, il vient

$$G(\mathbf{x}, 0) = F'(\mathbf{x}, 0) = f(\mathbf{x}) \quad \text{et} \quad G(\mathbf{x}, 1) = F_1(\mathbf{x}, 1) = g(\mathbf{x})$$

pour tout point $\mathbf{x} \in \mathscr{X}$, i.e. G est une homotopie reliant les applications f et g. En outre, G est une homotopie différentiable car l'application F' (donc l'application G) est différentiable pour $t < 3/7$ et $t > 4/7$, et l'application G est confondue avec P (donc est différentiable aussi) pour $2/7 < t < 5/7$.

Nous avons construit une homotopie différentiable, mais elle prend ses valeurs sur $\mathbb{R}^n$ et non pas sur la variété $\mathscr{Y}$. Pour remédier à cet état de fait on rappellera que la construction de l'homotopie G dépendait du paramètre $\varepsilon > 0$ et que

$$|G(\mathbf{x}, t) - F'(\mathbf{x}, t)| \leqslant \lambda(t)\,|P(\mathbf{x}, t) - F'(\mathbf{x}, t)| < \varepsilon$$

pour tout point $(\mathbf{x}, t) \in \mathscr{X} \times [0, 1]$. Puisque $F'(\mathbf{x}, t) \in \mathscr{Y}$, ceci exprime par définition que *la distance du point* $G(\mathbf{x}, t)$ *à la variété* $\mathscr{Y}$ *est* $< \varepsilon$, i.e. le point $G(\mathbf{x}, t)$ est contenu dans un ε-voisinage de la variété $\mathscr{Y}$.

D'autre part, il existe par hypothèse un voisinage $O_{\mathscr{Y}}$ de la variété $\mathscr{Y}$ appliqué sur $\mathscr{Y}$ par une rétraction. Ceci étant, puisque la variété $\mathscr{Y}$ est compacte, il existe un $\varepsilon > 0$ tel que le ε-voisinage de la variété $\mathscr{Y}$ est entièrement contenu dans $O_{\mathscr{Y}}$ (prouvez-le !). Donc, l'homotopie G construite pour ce ε sera telle que $G(\mathbf{x}, t) \in O_{\mathscr{Y}}$ pour tout

point $(\mathbf{x}, t) \in \mathscr{X} \times [0, 1]$. Donc, la formule

$$H(\mathbf{x}, t) = r_{\mathscr{Y}}(G(\mathbf{x}, t)), \quad (\mathbf{x}, t) \in \mathscr{X} \times [0, 1],$$

où $r_{\mathscr{Y}} : O_{\mathscr{Y}} \to \mathscr{Y}$ est une rétraction différentiable, définit une homotopie différentiable $H : \mathscr{X} \times [0, 1] \to \mathscr{Y}$ reliant l'application $r_{\mathscr{Y}} \circ f = f$ à l'application $r_{\mathscr{Y}} \circ g = g$.

Ceci prouve totalement la proposition 2 (sous l'hypothèse que les variétés $\mathscr{X}$ et $\mathscr{Y}$ sont des retracts de voisinage) et en même temps le théorème du tambour. □

Remarque 4. De même que nous avons approché une homotopie arbitraire $\mathscr{X} \times [0, 1] \to \mathscr{Y}$ par une homotopie différentiable, nous pouvons approcher toute application continue $f : \mathscr{X} \to \mathscr{Y}$ qar une application différentiable.

Exercice 2. Dans le prochain ouvrage on montrera que pour toute variété compacte $\mathscr{Y}$ plongée dans l'espace $\mathbf{R}^n$, il existe une constante $d > 0$ telle que deux points quelconques $p, q \in \mathscr{Y}$, dont la distance (mesurée sur $\mathscr{Y}$) soit $< d$, peuvent être reliés dans $\mathscr{Y}$ par un plus court chemin et un seul (pour la sphère, la constante d est égale à π, et le plus court chemin pour aller d'un point à un autre est l'arc de grand cercle reliant ces points). Utiliser ce fait pour prouver que *deux applications $\mathscr{X} \to \mathscr{Y}$ suffisamment voisines sont homotopes.*

On en déduit en particulier que deux applications différentiables $\mathscr{X} \to \mathscr{Y}$ approchant d'assez près une application continue donnée $f : \mathscr{X} \to \mathscr{Y}$ sont de même degré (si les variétés $\mathscr{X}$ et $\mathscr{Y}$ sont orientées, compactes et de même dimension). Ce degré s'appelle *degré de l'application continue f.*

Exercice 3. Montrer que *les applications continues homotopes sont de même degré.*

LEÇON 27

Domaines à frontière régulière. — Théorème de Stokes. — Formules de Gauss-Ostrogradski, Green et Newton-Leibniz. — Variétés à bord. — Points intérieurs et points du bord. — ∂-sous-variétés plongées. — Théorème de Stokes pour variétés à bord et ∂-sous-variétés. — Théorème de Stokes pour intégrales de surface. — Théorème de Stokes pour sous-variétés singulières. — Intégrales curvilignes de seconde espèce.

Supposons comme plus haut que $\mathscr{X}$ est une variété différentiable séparée et paracompacte à n dimensions.

Définition 1. On dit qu'un sous-ensemble D d'une variété $\mathscr{X}$ est un *domaine à frontière régulière* si

1° le sous-ensemble D est l'adhérence de son intérieur:

$$D = \overline{\operatorname{Int} D};$$

2° sa frontière

$$\operatorname{Fr} D = \bar{D} \setminus \operatorname{Int} D$$

est une sous-variété plongée à $n - 1$ dimensions.

La frontière $\operatorname{Fr} D$ d'un tel domaine D s'appelle *bord* et se note ∂D.

On sait (cf. leçon 13) que pour tout point de la sous-variété ∂D il existe dans la variété $\mathscr{X}$ une carte $(U, h) = (U, x^1, \ldots, x^n)$ contenant ce point, telle que

$$(V, k) = (V, y^1, \ldots, y^{n-1}),$$

où $V = U \cap \partial D$ et $y^1 = x^2 |_V, \ldots, y^{n-1} = x^n |_V$, est une carte sur ∂D, et l'égalité $x^1(p) = 0$, $p \in U$, a lieu si et seulement si $p \in \partial D$. Ceci étant, sans perdre en généralité on peut supposer que $x^1 < 0$ sur $U \cap \operatorname{Int} D$.

On dira de la carte (U, h) possédant ces propriétés qu'elle est *adaptée à* D, et de la carte (V, k) qu'elle est *découpée sur* ∂D par la carte (U, h).

Soient maintenant $(U, h) = (U, x^1, \ldots, x^n)$ et $(U', h') = (U', x^{1'}, \ldots, x^{n'})$ deux cartes sur $\mathscr{X}$ adaptées à D, et (V, k), (V', k') les cartes qu'elles découpent sur ∂D. La fonction $x^{1'}$ étant identiquement nulle sur $V \cap V'$, on a sur $V \cap V'$

$$\frac{\partial x^{1'}}{\partial x^k} = 0, \quad k = 2, \ldots, n.$$

Donc,

$$\det \frac{\partial h'}{\partial h} = \frac{\partial x^{1'}}{\partial x^1} \cdot \det \frac{\partial k'}{\partial k} \quad \text{sur } V \cap V'.$$

D'autre part, puisque $x^{1'} < 0$ si et seulement si $x^1 < 0$, il vient

$$\frac{\partial x^{1'}}{\partial x^1} > 0 \quad \text{sur } V \cap V'.$$

Donc, *si les cartes* (U, h) *et* (U', h') *sont positivement compatibles, il en est de même des cartes* (V, k) *et* (V', k') *(pour* $n > 1$*).*

Supposons maintenant que la variété $\mathscr{X}$ est orientable et orientée. Etant donné que pour $n > 1$ tout point p de ∂D est visiblement contenu dans une carte positive adaptée à D, toutes les cartes positives adaptées à D découpent sur ∂D un atlas de cartes positivement compatibles. On dira de l'orientation définie sur ∂D par cet atlas qu'elle est *induite* par celle de la variété $\mathscr{X}$.

Pour $n = 1$ le domaine D est un système de segments (sur une droite ou un cercle) et ∂D est composé de leurs extrémités. L'orientation de la variété $\mathscr{X}$ induit un sens sur ces segments et l'on munit ∂D d'une orientation (au sens de la remarque 1 de la leçon 25) en admettant que l'extrémité droite de chaque segment est affectée du signe $+$ et l'extrémité gauche du signe $-$.

Donc, *dans une variété orientable (resp. orientée)* $\mathscr{X}$ *le bord de tout domaine à frontière régulière est une variété orientable (resp. orientée).*

* * *

Puisqu'en vertu du théorème de Sard (cf. leçon 15) le bord d'un domaine arbitraire à frontière régulière est un ensemble négligeable, tout domaine compact D à frontière régulière est cubable. Donc, si l'on admet que la variété $\mathscr{X}$ est orientée, l'intégrale

$$\int_D \omega$$

est définie pour toute forme ω de degré n sur $\mathscr{X}$ et tout domaine compact à frontière régulière. Cette intégrale est définie aussi pour les domaines D non compacts si seulement la forme ω est à support borné.

En particulier, pour toute forme ω de degré $n - 1$ (à support borné si le domaine D n'est pas compact) est définie l'intégrale

$$\int_D d\omega. \tag{1}$$

Par ailleurs, est définie aussi (par rapport à l'orientation induite sur la variété ∂D) l'intégrale

$$\int_{\partial D} i^*\omega,$$

où $i: \partial D \to \mathscr{X}$ est une injection. Pour simplifier les formules, on désignera cette intégrale par le symbole

$$\int_{\partial D} \omega. \tag{2}$$

Théorème 1 (théorème de Stokes pour domaines à frontières régulières). *Pour tout domaine D à frontière régulière d'une variété $\mathscr{X}$ séparée paracompacte orientée à n dimensions et toute forme $\omega \in \Omega^{n-1}\mathscr{X}$ (à support borné si le domaine D n'est pas compact) on a l'égalité*

$$\int_D d\omega = \int_{\partial D} \omega. \tag{3}$$

D é m o n s t r a t i o n. Il est clair que toutes les cartes (U_α, h_α) de la variété $\mathscr{X}$, qui soit ne rencontrent pas ∂D, soit sont adaptées à D, forment un atlas sur $\mathscr{X}$. La variété $\mathscr{X}$ étant par hypothèse paracompacte et séparée, il existe un recouvrement $\{U_\alpha\}$ subordonné à une partition $\{\eta_\alpha\}$ de l'unité. Puisque $\omega = \sum_\alpha \eta_\alpha \omega$ et $d\omega = \sum_\alpha d(\eta_\alpha \omega)$, on a

$$\int_D d\omega = \sum_\alpha \int_D d(\eta_\alpha \omega)$$

et

$$\int_{\partial D} \omega = \sum_\alpha \int_{\partial D} \eta_\alpha \omega$$

(il est clair que l'intégrale est additive aussi par rapport aux sommes infinies envisagées dont seul un nombre fini de termes sont non nuls au voisinage de tout point). Donc, il suffit de prouver la formule (3) pour une forme $\eta_\alpha \omega$, i.e. sous l'hypothèse que $\omega = 0$ à l'extérieur d'une carte positive $(U, h) = (U, x^1, \ldots, x^n)$ de la variété $\mathscr{X}$, qui soit ne rencontre pas ∂D, soit est adaptée à D. Ceci étant, si sur U

$$\omega = \sum_{k=1}^{n} (-1)^{k-1} w_k \, dx^1 \wedge \ldots \wedge \widehat{dx^k} \wedge \ldots \wedge dx^n,$$

alors

$$d\omega = \left(\sum_{k=1}^{n} \frac{\partial w_k}{\partial x^k} \right) dx^1 \wedge \ldots \wedge dx^n \quad \text{sur } U$$

(cf. formule (4) de la leçon 19) et par suite

$$\int_D d\omega = \int_{U \cap D} d\omega = \int_{h(U \cap D)} \left(\sum_{k=1}^{n} \frac{\partial w_k}{\partial x^k} \right) dx^1 \ldots dx^n.$$

Les fonctions $w_1, \ldots, w_n$ traitées comme des fonctions sur $h(U)$ sont nulles à l'extérieur d'un ensemble fermé, et si on les prolonge par 0 à l'extérieur de $h(U)$ à $\mathbb{R}^n$ tout entier, on obtient encore des fonctions différentiables. D'autre part, on peut admettre sans perdre en généralité que l'ensemble ouvert $h(U)$ est borné, i.e. contenu dans un cube

$$I_R^n = \{x \in \mathbb{R}^n, \ |x^1| \leqslant R, \ldots, |x^n| \leqslant R\}$$

(sur la frontière duquel les fonctions w_k sont donc toutes nulles).

Cas 1. La carte (U, h) ne rencontre pas ∂D. Dans ce cas l'intégrale (2) est visiblement nulle (car $i^*\omega = 0$), et par conséquent, pour prouver la formule (3) il suffit de montrer que l'intégrale (1) est nulle. Ceci étant, on peut admettre sans perdre en généralité que soit $U \subset \mathscr{X} \setminus D$, soit $U \subset D$. Mais pour $U \subset \mathscr{X} \setminus D$ l'intégrale (1) est visiblement nulle, et pour $U \subset D$ elle s'exprime par la formule

$$\int_D d\omega = \int_{I_R^n} \left(\sum_{k=1}^n \frac{\partial w_k}{\partial x^k} \right) dx^1 \ldots dx^n = \sum_{k=1}^n \int_{-R}^R \ldots \int_{-R}^R \frac{\partial w_k}{\partial x^k} dx^1 \ldots dx^n$$

qui est nulle puisque l'intégrale

$$\int_{-R}^R \frac{\partial w_k}{\partial x^k} dx^k = w_k(x^1, \ldots, R, \ldots, x^n) - w_k(x^1, \ldots, -R, \ldots, x^n)$$

l'est pour tout $k = 1, \ldots, n$.

Cas 2. La carte (U, h) est adaptée à D. On a alors par analogie

$$\int_D d\omega = \sum_{k=1}^n \int_{-R}^0 \int_{-R}^R \ldots \int_{-R}^R \frac{\partial w_k}{\partial x^k} dx^1 \ldots dx^k =$$

$$= \int_{-R}^R \ldots \int_{-R}^R w_1(0, x^2, \ldots, x^n) dx^2 \ldots dx^n.$$

D'autre part, puisque $i^*(dx^1) = 0$ et $i^*(dx^2) = dy^1, \ldots, i^*(dx^n) = dy^{n-1}$ (car $x^1 = 0$ et $x^2 = y^1, \ldots, x^n = y^{n-1}$ sur ∂D), alors

$$i^*\omega = w_1(0, y^1, \ldots, y^{n-1}) dy^1 \wedge \ldots \wedge dy^{n-1} \quad \text{sur } U \cap \partial D.$$

Donc,

$$\int_{-R}^R \ldots \int_{-R}^R w_1(0, x^2, \ldots, x^n) dx^2 \ldots dx^n =$$

$$= \int_{-R}^R \ldots \int_{-R}^R w_1(0, y^1, \ldots, y^{n-1}) dy^1 \ldots dy^{n-1} = \int_{U \cap \partial D} i^*\omega = \int_{\partial D} \omega$$

Par conséquent,

$$\int_D d\omega = \int_{\partial D} \omega. \quad \square$$

Corollaire 1. *Pour toute forme ω à support borné de degré $n-1$ sur une variété $\mathscr{X}$ séparée paracompacte orientée à n dimensions on a l'égalité*

$$\int_{\mathscr{X}} d\omega = 0. \tag{4}$$

Démonstration. Soient p_0 un point de la variété $\mathscr{X}$, et (U, h) une carte de $\mathscr{X}$ telle que $p_0 \in U$ et $h(U) = \mathring{\mathbb{B}}_2^n$, où comme toujours $\mathring{\mathbb{B}}_2^n$ est une boule ouverte de $\mathbb{R}^n$ de rayon 2 et de centre 0. Supposons ensuite que $D_1 = h^{-1}(\mathbb{B}_1^n)$, où $\mathbb{B}_1^n$ est une boule fermée concentrique de rayon 1, et soit $D_2 = \mathscr{X} \setminus \operatorname{Int} D_1$. Il est clair que les ensembles D_1 et D_2 sont des domaines possédant une frontière régulière commune $h^{-1}(\mathbb{S}^{n-1})$. (On dit que D_2 s'obtient à partir de $\mathscr{X}$ par *forage de la boule* D_1.) Ceci étant, les orientations induites sur le bord $h^{-1}(\mathbb{S}^{n-1})$ par celle de la variété $\mathscr{X}$ sont visiblement contraires (ce qu'on peut exprimer par la formule $\partial D_2 = -\partial D_1$), et donc,

$$\int_{\partial D_2} \omega = -\int_{\partial D_1} \omega.$$

En appliquant le théorème 1 aux domaines D_1 et D_2 et en tenant compte du fait que

$$\int_{\mathscr{X}} d\omega = \int_{D_1} d\omega + \int_{D_2} d\omega,$$

on obtient immédiatement (4). $\square$

La formule (4) peut être traitée comme un cas particulier de la formule générale (3) si l'on convient que l'intégrale d'une forme arbitraire étendue à un ensemble vide est nulle.

Nous nous sommes déjà servis du corollaire 1 à la leçon 25 (cf. page 366).

* * *

Dans le cas particulier où $\mathscr{X}$ est l'espace $\mathbb{R}^3$ muni des coordonnées x, y, z, chaque forme ω de degré 3 sur $\mathscr{X}$ se représente par $f\, dx \wedge dy \wedge dz$, où f est une fonction, et pour tout ensemble cubable $D \subset \mathbb{R}^3$ l'intégrale $\int_D \omega$ est égale à l'intégrale de Riemann $\int_D f(\mathbf{x})\, d\mathbf{x}$

qui dans le cas envisagé est généralement désignée par le symbole

$$\iiint_D f \, dx \, dy \, dz$$

pour souligner la triple dimension et la possibilité de passer à une intégrale triple de Riemann sur une droite. Par analogie, pour toute surface orientée $\mathscr{X}$ de $\mathbb{R}^3$ (qui est une sous-variété orientée à deux dimensions) et toute forme

$$\omega = P \, dy \wedge dz + Q \, dz \wedge dx + R \, dx \wedge dy$$

sur $\mathbb{R}^3$ (nous attirons l'attention du lecteur sur l'ordre des différentielles du deuxième terme de la somme), l'intégrale $\int_{\mathscr{X}} \omega$ (plus exactement, l'intégrale $\int_{\mathscr{X}} i^* \omega$, où i est une injection de $\mathscr{X}$ dans $\mathbb{R}^3$) est désignée par le symbole

$$\iint_{\mathscr{X}} P \, dy \, dz + Q \, dz \, dx + R \, dx \, dz. \tag{5}$$

Puisque

$$d\omega = \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dx \wedge dy \wedge dz,$$

on obtient le cas particulier suivant du théorème 1 : *pour tout domaine $D \subset \mathbb{R}^3$ à frontière régulière et toutes fonctions P, Q et R on a la formule*

$$\iiint_D \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dx \, dy \, dz = \tag{6}$$

$$= \iint_{\partial D} P \, dy \, dz + Q \, dz \, dx + R \, dx \, dy.$$

La formule (6) s'appelle *formule de Gauss-Ostrogradski.*

Remarque 1. Il est évident que dans la formule (6) les intégrales sont toutes supposées exister, i.e. soit les fonctions P, Q et R sont à support borné, soit le domaine D est compact (et les fonctions P, Q et R sont évidemment définies sur D). A noter que, pour que cette formule soit valable, il n'est pas nécessaire que D soit un domaine à frontière régulière, il suffit, disons, de supposer qu'il soit à frontière *régulière par morceaux* (dans un sens connu).

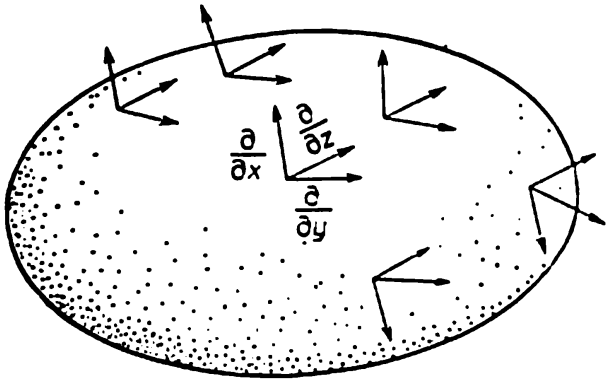
Remarque 2. L'orientation d'un plan de $\mathbb{R}^3$ est définie par l'une de ses faces, i.e. par un vecteur orthogonal à ce plan. Donc, l'orientation d'une surface de $\mathbb{R}^3$ est définie par un champ de vecteurs normaux non nuls. Il est aisé de voir que ce champ est composé de *nor-*

males extérieures pour l'orientation induite sur le bord ∂D dans la formule (6).

Sur le plan $\mathbb{R}^2$ l'analogue de la formule (6) est

$$\iint_D \left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right) \mathrm{d}x\,\mathrm{d}y=\int_{\partial D} P\,\mathrm{d}x+Q\,\mathrm{d}y \tag{7}$$

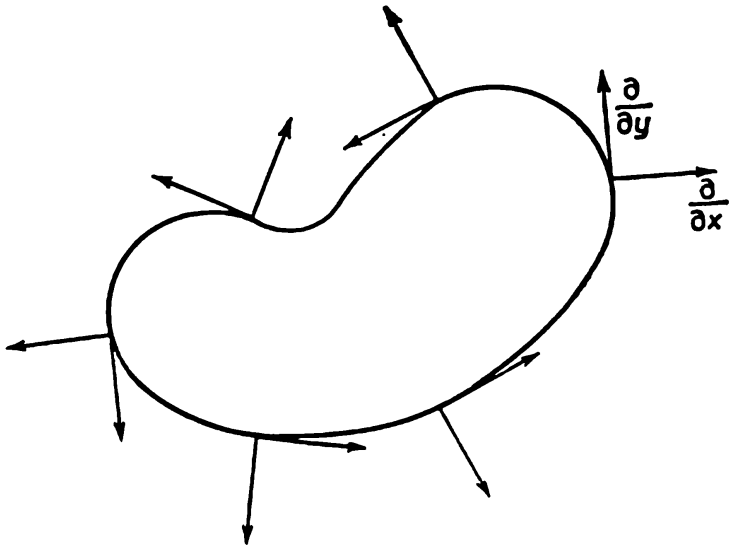
et s'appelle *formule de Green*. L'orientation de ∂D, i.e. le sens de parcours, est telle que le domaine D doit rester à gauche.



Enfin sur la droite $\mathbb{R}$ la formule (3) se transforme en la *formule de Newton-Leibniz*

$$\int_a^b f'(x)\,\mathrm{d}x=f(b)-f(a). \tag{8}$$

(On rappelle qu'une intégrale sur une variété orientée 0-dimensionnelle est définie par la formule (3) de la leçon 25 et que par définition les points b et a du bord de $[a, b]$ sont affectés respectivement du signe $+$ et du signe $-$.)



* * *

La notion de domaine à frontière régulière et le théorème 1 admettent une importante généralisation.

Soit $\mathbb{R}^n_{(-)}$ le demi-espace de $\mathbb{R}^n$ composé des points $\mathbf{x} = (x^1, \ldots, x^n)$ tels que $x^1 \leqslant 0$ et soit $\mathbb{R}^{n-1}_0$ son *bord* qui est constitué des points $\mathbf{x}$ tels que $x^1 = 0$.

On dira qu'un sous-ensemble $U \subset \mathbb{R}^n$ est *∂-ouvert* s'il est ouvert soit dans $\mathbb{R}^n$, soit dans $\mathbb{R}^n_{(-)}$, i.e. s'il existe un ensemble ouvert U' dans $\mathbb{R}^n$ tel que soit $U = U'$, soit $U = U' \cap \mathbb{R}^n_{(-)}$. On dira qu'une application $\varphi: U \to \mathbb{R}^m$ définie sur un ensemble ∂-ouvert $U \subset \mathbb{R}^n$ est *différentiable* s'il existe un ensemble ouvert $U' \subset \mathbb{R}^n$ et une application différentiable $\varphi': U' \to \mathbb{R}^m$ tels que $U \subset U'$ et $\varphi' \mid_U = \varphi$. Ceci équivaut à l'existence des dérivées partielles continues de tous les ordres nécessaires des fonctions définissant l'application φ (à condition, lorsque $U \subset \mathbb{R}^n_{(-)}$ et $U \cap \mathbb{R}^{n-1}_0 \neq \varnothing$, que la dérivation par rapport à x^1 soit comprise comme une dérivation à d r o i t e en tout point de $U \cap \mathbb{R}^{n-1}_0$). Une application $\varphi: U \to V$ d'ensembles ∂-ouverts s'appelle *difféomorphisme* si elle est bijective, différentiable et si sa réciproque $\varphi^{-1}: V \to U$ est différentiable aussi.

Soit $\mathscr{D}$ un ensemble quelconque. On appellera *∂-carte* dans $\mathscr{D}$ un couple (U, h) composé d'un sous-ensemble $U \subset \mathscr{D}$ et d'une application bijective $h: U \to h(U)$ sur l'ensemble ∂-ouvert $h(U) \subset \mathbb{R}^n$. On dira que deux ∂-cartes (U, h) et (V, k) sont *compatibles* si soit $U \cap V = \varnothing$, soit $U \cap V \neq \varnothing$ et

a) les ensembles $h(U \cap V)$ et $k(U \cap V)$ sont des sous-ensembles ∂-ouverts de l'espace $\mathbb{R}^n$;

b) l'application

$$k \circ h^{-1}: h(U \cap V) \to k(U \cap V)$$

est un difféomorphisme. Il est clair que toute carte de $\mathscr{D}$ (au sens de la définition 1 de la leçon 6) est une ∂-carte et que les cartes compatibles (au sens de la définition 2 de la leçon 6) le sont en tant que ∂-cartes.

On appelle *atlas de ∂-cartes* sur l'ensemble $\mathscr{D}$ (cf. définition 3 de la leçon 6) un ensemble de ∂-cartes (U_α, h_α) deux à deux compatibles et telles que leurs supports U_α forment un recouvrement de $\mathscr{D}$. Il est aisé de voir (cf. proposition 1 de la leçon 6 et son corollaire 1) que tout atlas de ∂-cartes A est contenu dans l'unique atlas maximal composé de toutes les ∂-cartes compatibles avec les ∂-cartes de A.

Définition 2. Un ensemble $\mathscr{D}$ dans lequel est défini un atlas maximal de ∂-cartes s'appelle *∂-variété différentiable* (et ses cartes, *cartes différentiables*).

Il est évident que toute variété $\mathscr{X}$ est une ∂-variété.

Soulignons que comme à la leçon 6 on admet que la dimension n de l'espace $\mathbb{R}^n$ est fixée (cette dimension s'appelle *dimension* de la ∂-variété $\mathscr{D}$ et se note $\dim \mathscr{D}$), de même que la classe C^r des appli-

cations envisagées, où r est un entier positif ou bien l'un des symboles ∞ ou ω. (Cette classe s'appelle *classe de différentiabilité* de la ∂-variété $\mathscr{D}$; on admettra généralement que $r = \infty$.)

On munit une ∂-variété d'une topologie de la même façon qu'une variété, i.e. un sous-ensemble O d'une ∂-variété $\mathscr{D}$ est ouvert si et seulement si pour toute carte différentiable (U, h) l'ensemble $h(O \cap U)$ est ∂-ouvert dans $\mathbb{R}^n$. Donc, toute ∂-variété est un espace topologique (qui satisfait au premier axiome de dénombrabilité et qui est de plus localement compact si la ∂-variété est séparée).

* * *

Définition 3. On appelle *carte intérieure* d'une ∂-variété $\mathscr{D}$ une ∂-carte (U, h) pour laquelle l'ensemble $h(U)$ est ouvert dans $\mathbb{R}^n$. Un point $p \in \mathscr{D}$ s'appelle *point intérieur* d'une ∂-variété $\mathscr{D}$ si dans $\mathscr{D}$ il existe une carte intérieure (U, h) telle que $p \in U$. L'ensemble de tous les points intérieurs d'une ∂-variété $\mathscr{D}$ est désigné par $\operatorname{int} \mathscr{D}$ et s'appelle *intérieur* de $\mathscr{D}$.

Il est clair que $\operatorname{int} \mathscr{D}$ est ouvert dans $\mathscr{D}$ et est une variété différentiable (avec un atlas composé de toutes les cartes intérieures).

Exercice 1. Montrer que $\operatorname{int} \mathscr{D}$ n'est pas vide (si $\mathscr{D}$ ne l'est pas) et de plus que

$$\overline{\operatorname{int} \mathscr{D}} = \mathscr{D}.$$

Définition 4. On appelle *carte de bord* d'une ∂-variété $\mathscr{D}$ une ∂-carte (U, h) telle que $h(U) \subset \mathbb{R}^n_{(-)}$ et $h(U) \cap \mathbb{R}^{n-1}_0 \neq \varnothing$. Un point $p \in \mathscr{D}$ s'appelle *point du bord* d'une ∂-variété $\mathscr{D}$ s'il existe dans $\mathscr{D}$ une carte de bord (U, h) telle que $p \in U$ et $h(p) \in h(U) \cap \mathbb{R}^{n-1}_0$. L'ensemble (éventuellement vide) de tous les points du bord est désigné par $\partial \mathscr{D}$ et appelé *bord* de la ∂-variété $\mathscr{D}$. Par définition,

$$\mathscr{D} = \partial \mathscr{D} \cup \operatorname{int} \mathscr{D}.$$

Proposition 1. *Aucun point intérieur n'est point du bord, et réciproquement, aucun point du bord n'est point intérieur*:

$$\partial \mathscr{D} \cap \operatorname{int} \mathscr{D} = \varnothing.$$

D é m o n s t r a t i o n. Affirmer que $p \in \partial \mathscr{D} \cap \operatorname{int} \mathscr{D}$ revient à affirmer qu'il existe dans $\mathscr{D}$ un ensemble ouvert U contenant le point p, et des applications $h : U \to \mathbb{R}^n_{(-)}$ et $k : U \to \mathbb{R}^n$, tels que le couple (U, h) est une carte de bord et le couple (U, k) une carte intérieure de $\mathscr{D}$. Donc, l'ensemble $k(U)$ est ouvert dans $\mathbb{R}^n$ mais pas $h(U)$. L'application $h \circ k^{-1}$ étant un difféomorphisme de l'ensemble $k(U)$ sur l'ensemble $h(U)$, ceci contredit le lemme 1 établi plus bas. Donc, le point p n'existe pas, et par suite, $\partial \mathscr{D} \cap \operatorname{int} \mathscr{D} = \varnothing$. □

Lemme 1. *Soit* $\varphi : O \to \mathbb{R}^n$ *une application différentiable (de classe* C^r, $r \geqslant 1$) *d'un ensemble ouvert* $O \subset \mathbb{R}^n$ *dans* $\mathbb{R}^n$. *Si le jacobien*

D_φ de l'application φ est non nul en tout point x *de O, l'ensemble $\varphi(O)$ est ouvert.*

Démonstration. D'après le théorème de l'application réciproque (cf. leçon 6) le point $x \in O$ possède un voisinage $U \subset O$ s'appliquant sur un voisinage V du point $\varphi(x)$. Comme $V \subset \varphi(O)$, on a donc $\varphi(x) \in \operatorname{Int} \varphi(O)$, et puisque ceci est valable pour tout point $x \in O$, il vient $\varphi(O) = \operatorname{Int} \varphi(O)$, i.e. l'ensemble $\varphi(O)$ est ouvert.□

La proposition 1 est valable aussi pour $r = 0$ (pour les variétés topologiques). L'analogue du lemme 1 (connu sous le nom de théorème de Brouwer d'invariance d'un domaine) affirme que si une application continue $\varphi: O \to \mathbb{R}^n$ d'un ensemble ouvert $O \subset \mathbb{R}^n$ dans $\mathbb{R}^n$ est un monéomorphisme (un homéomorphisme sur $\varphi(O)$), l'ensemble $\varphi(O)$ est ouvert. Nous manquons malheureusement de place pour produire la démonstration assez longue et fastidieuse de cette assertion.

Corollaire 1. *Le bord $\partial\mathscr{D}$ d'une ∂-variété $\mathscr{D}$ est fermé dans $\mathscr{D}$.*

Démonstration. En vertu de la proposition 1 on a $\partial\mathscr{D} = \mathscr{D} \setminus \operatorname{int} \mathscr{D}$. Or $\operatorname{int} \mathscr{D}$ est ouvert dans $\mathscr{D}$.□

Corollaire 2. *L'égalité $\partial\mathscr{D} = \varnothing$ a un sens si et seulement si la ∂-variété $\mathscr{D}$ est une variété.*

Démonstration. Il est clair que $\mathscr{D}$ est une variété si et seulement si $\operatorname{int} \mathscr{D} = \mathscr{D}$.□

D'après le corollaire 2 les variétés différentiables s'appellent aussi *variétés sans bord*. Par voie de conséquence on appellera *variété à bord* une ∂-variété telle que $\partial\mathscr{D} \neq \varnothing$. Du reste, le terme « variété à bord » est souvent utilisé comme synonyme de «∂-variété ». (Lorsqu'il y a risque d'ambiguïté, on parle des « variétés à bord non vide » ou respectivement des « variétés à ou sans bord ».)

Les variétés compactes sans bord sont d'une importance toute particulière. On les appelle *variétés fermées*.

Pour toute carte de bord $(U, h) = (U, x^1, \ldots, x^n)$ d'une ∂-variété $\mathscr{D}$ le couple $(U \cap \partial\mathscr{D}, h_{U \cap \partial\mathscr{D}})$ est en raison de l'identification $\mathbb{R}_0^{n-1} = \mathbb{R}^{n-1}$ une carte dans $\partial\mathscr{D}$ (avec les coordonnées locales $x^2, \ldots, x^n$), et de telles cartes forment un atlas sur $\partial\mathscr{D}$. Ceci montre que *le bord $\partial\mathscr{D}$ d'une ∂-variété $\mathscr{D}$ à n dimensions est une variété sans bord à $n - 1$ dimensions.*

* * *

Il est clair que *tout domaine D à frontière régulière est une ∂-variété à bord ∂D*. (Les ∂-cartes de cette ∂-variété sont soit des cartes de la variété ambiante $\mathscr{X}$ contenues dans $\operatorname{Int} D$, soit les intersections des cartes adaptées à D avec D.) La dimension de cette ∂-variété est égale à la dimension n de la variété $\mathscr{X}$.

Une application $f: \mathscr{D} \to \mathscr{X}$ d'une ∂-variété $\mathscr{D}$ dans une ∂-variété $\mathscr{X}$ (nous n'aurons besoin que du cas où $\mathscr{X}$ est sans bord) est dite

différentiable si elle est continue et si pour deux ∂-cartes quelconques (U, h) dans $\mathscr{D}$ et (V, k) dans $\mathscr{X}$, telles que $fU \subset V$, l'application composée

$$k \circ f \circ h^{-1} : h\,(U) \to k\,(V)$$

est différentiable.

Supposons que la variété $\mathscr{X}$ est sans bord (le cas où $\partial\mathscr{X} \neq \varnothing$ apporte des complications sur lesquelles nous glisserons) et que $\mathscr{D} \subset \mathscr{X}$. Si l'injection $i : \mathscr{D} \to \mathscr{X}$ est différentiable et les variétés int $\mathscr{D}$ et $\partial\mathscr{D}$ sont des sous-variétés plongées dans $\mathscr{X}$, alors $\mathscr{D}$ s'appelle *∂-sous-variété* (*plongée*) de la variété $\mathscr{X}$.

Par exemple, tout domaine à frontière régulière est une ∂-sous-variété.

Exercice 2. Montrer que si une ∂-sous-variété $\mathscr{D}$ d'une variété $\mathscr{X}$ est de même dimension que $\mathscr{X}$, alors $\mathscr{D}$ est un domaine à frontière régulière dans $\mathscr{X}$.

On appellera *forme différentielle* ω *de degré* m sur une ∂-variété $\mathscr{D}$ une forme sur int $\mathscr{D}$ telle que pour toute carte de bord (U, h) les coefficients $\omega_{i_1 \ldots i_m}$ de cette forme dans la carte $(U \cap \operatorname{int} \mathscr{D}, h|_{U \cap \operatorname{int}\mathscr{D}})$ de la variété int $\mathscr{D}$ sont les restrictions de certaines fonctions différentiables (visiblement définies de façon unique) sur U. Les dernières fonctions seront désignées par les mêmes symboles $\omega_{i_1 \ldots i_m}$ et appelées *coefficients de la forme* ω *dans la carte* (U, h). Les restrictions à $U \cap \partial\mathscr{D}$ des coefficients $\omega_{i_1 \ldots i_m}$, où $i_1, \ldots, i_m \neq 1$, sont de toute évidence les coefficients d'une forme de degré $m - 1$ sur $\partial\mathscr{D}$ que l'on désignera par $\omega|_{\partial\mathscr{D}}$.

Lorsque $\mathscr{D}$ est une ∂-sous-variété d'une variété $\mathscr{X}$ sans bord (par exemple, un domaine à frontière régulière), pour toute forme ω sur $\mathscr{X}$ la forme $j^*\omega$, où $j : \operatorname{int} \mathscr{D} \to \mathscr{X}$ est une injection, est visiblement une forme sur $\mathscr{D}$. Pour cette forme, $j^*\omega|_{\partial\mathscr{D}} = i^*\omega$, où i est une injection de $\partial\mathscr{D}$ dans $\mathscr{X}$.

Exercice 3. Montrer que toute forme sur $\mathscr{D}$ se représente par $j^*\omega$, où ω est une forme sur $\mathscr{X}$.

On dit qu'une variété à bord $\mathscr{D}$ est *orientable* (resp. *orientée*) s'il en est de même de la variété int $\mathscr{D}$. Une carte de bord (U, h) d'une ∂-variété orientée $\mathscr{D}$ est dite *positive* si la carte intérieure $(U \cap \operatorname{int} \mathscr{D},\ h|_{U \cap \operatorname{int}\mathscr{D}})$ l'est.

* * *

On dira qu'une forme différentielle sur une ∂-variété $\mathscr{D}$ est *à support borné* si elle est nulle à l'extérieur d'un ensemble compact $C \subset \mathscr{D}$. Si une ∂-variété $\mathscr{D}$ séparée et paracompacte est orientée, pour toute forme ω à support borné de degré $n = \dim \mathscr{D}$ est définie

l'intégrale

$$\int_{\operatorname{int}\mathscr{D}} \omega.$$

On appellera cette intégrale *intégrale de* ω *sur* $\mathscr{D}$ et on la désignera par

$$\int_{\mathscr{D}} \omega. \tag{9}$$

Si $\mathscr{D}$ est un domaine D à frontière régulière dans une variété $\mathscr{X}$, et ω, la restriction d'une forme ω sur $\mathscr{X}$ (à support borné si D n'est pas compact), l'intégrale (9) est égale à l'intégrale

$$\int_{D} \omega'$$

puisque le bord ∂D est un ensemble négligeable en vertu du théorème de Sard.

La construction de l'orientation induite sur le bord d'un domaine à frontière régulière s'étend *ad litteram* à toutes ∂-variétés orientées. Dans la suite on conviendra que le bord de toute ∂-variété orientée $\mathscr{D}$ est muni d'une orientation induite.

En particulier, ceci nous permet d'envisager l'intégrale

$$\int_{\partial\mathscr{D}} \omega\,|_{\partial\mathscr{D}}$$

pour toute forme ω à support borné de degré $n-1$ sur une ∂-variété orientée séparée et paracompacte à n dimensions, intégrale que nous écrivons brièvement

$$\int_{\partial\mathscr{D}} \omega. \tag{10}$$

Théorème 1′ (théorème de Stokes pour variétés à bord). *Pour toute forme* ω *à support borné de degré* $n-1$ *sur une* ∂*-variété* $\mathscr{D}$ *séparée paracompacte orientée à* n *dimensions, on a la formule*

$$\int_{\mathscr{D}} d\omega = \int_{\partial\mathscr{D}} \omega. \tag{11}$$

D é m o n s t r a t i o n. Elle reprend littéralement celle du théorème 1. Nous la laissons au soin du lecteur. □

On conviendra que l'intégrale (10) est nulle pour $\partial\mathscr{D} = \varnothing$. Aux termes de cette convention, le corollaire 1 du théorème 1 est un cas particulier du théorème 1′.

Lorsque $\mathscr{D}$ est une ∂-sous-variété d'une variété $\mathscr{X}$, pour toute forme ω de degré n sur $\mathscr{X}$ (on remarquera que n est la dimension de $\mathscr{D}$ et non pas de $\mathscr{X}$!) telle que $j^*\omega$ est à support borné sur $\mathscr{D}$ (j est une injection de $\mathscr{D}$ dans $\mathscr{X}$) on désignera l'intégrale

$$\int_{\mathscr{D}} j^*\omega$$

par

$$\int_{\mathscr{D}} \omega \tag{12}$$

(les formes ω sur $\mathscr{X}$ pour lesquelles la forme $j^*\omega$ est à support borné seront appelées *formes à support borné sur* $\mathscr{D}$).

Aux termes de cette convention on a de toute évidence la généralisation suivante du théorème 1 :

Théorème 1″ (théorème de Stokes pour ∂-sous-variétés). *Pour toute ∂-sous-variété orientée $\mathscr{D}$ à n dimensions d'une variété séparée et paracompacte $\mathscr{X}$ et toute forme ω à support borné sur $\mathscr{D}$ de degré $n-1$, on a la formule* (11) *sur* $\mathscr{X}$. □

* * *

Les ∂-sous-variétés à deux dimensions de l'espace $\mathbb{R}^3$ s'appellent *surfaces à bord*. Pour une telle surface $\mathscr{D}$ et une forme arbitraire

$$\omega = P\,\mathrm{d}y \wedge \mathrm{d}z + Q\,\mathrm{d}z \wedge \mathrm{d}x + R\,\mathrm{d}x \wedge \mathrm{d}y$$

sur $\mathbb{R}^3$ l'intégrale (12) est désignée par

$$\iint_{\mathscr{D}} P\,\mathrm{d}y\,\mathrm{d}z + Q\,\mathrm{d}z\,\mathrm{d}x + R\,\mathrm{d}x\,\mathrm{d}y \tag{13}$$

(comparer avec (5)). La formule (11) (pour la forme $\omega = P\,\mathrm{d}x + Q\,\mathrm{d}y + R\,\mathrm{d}z$) devient alors

$$\iint_{\mathscr{D}} \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathrm{d}y\,\mathrm{d}z + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathrm{d}z\,\mathrm{d}x + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathrm{d}x\,\mathrm{d}y = \int_{\partial\mathscr{D}} P\,\mathrm{d}x + Q\,\mathrm{d}y + R\,\mathrm{d}z. \tag{14}$$

La formule (14) est connue sous le nom de *formule de Stokes pour intégrales de surface dans* $\mathbb{R}^3$.

Remarque 3. Il est intéressant de noter pour la petite histoire que le nom de « formule de Stokes » a été attribué d'abord à la formule (14) et ensuite à ses généralisations et analogues (3) et (11)

à la suite d'une ambiguïté. La formule (14) était connue au milieu du XIXe siècle à l'Université de Cambridge et, sur proposition du célèbre physicien et mathématicien Thomson (qui en est probablement l'auteur), fut incluse dans les billets d'examen de mathématiques. Elle fut attribuée — d'abord par les étudiants seuls — à Stokes uniquement parce qu'il était président du jury.

* * *

Les théorèmes 1′ et 1″ peuvent être regroupés en un seul théorème général.

Soient $\mathscr{D}$ et $\mathscr{X}$ des ∂-variétés et soit $\gamma : \mathscr{D} \to \mathscr{X}$ une application différentiable de int $\mathscr{D}$ dans int $\mathscr{X}$. Alors pour toute forme ω sur $\mathscr{X}$ la forme $\gamma^*\omega$ (définie sur int $\mathscr{D}$) sera de toute évidence une forme sur $\mathscr{D}$. Donc, si elle est à support borné et de degré $n = \dim \mathscr{D}$, et si la variété $\mathscr{D}$ est orientable, alors est définie l'intégrale

$$\int_{\mathscr{D}} \gamma^*\omega. \tag{15}$$

Les applications différentiables $\gamma : \mathscr{D} \to \mathscr{X}$ de int $\mathscr{D}$ dans int $\mathscr{X}$ pour lesquelles la ∂-variété $\mathscr{D}$ est orientée et à n dimensions seront appelées *sous-variétés singulières à n dimensions* de la ∂-variété $\mathscr{X}$, les formes ω sur $\mathscr{X}$ pour lesquelles la forme $\gamma^*\omega$ est à support borné, *formes à support borné sur* γ, et enfin l'intégrale (15), *intégrale de* ω *sur* γ. D'après ce qui précède on désignera l'intégrale (15) par

$$\int_{\gamma} \omega. \tag{16}$$

On admettra aussi que l'« application » $\varnothing \to \mathscr{X}$ de l'ensemble vide $\varnothing$ dans $\mathscr{X}$ est une sous-variété singulière à n dimensions et que l'intégrale (16) étendue à une telle sous-variété est nulle.

Aux termes de cette convention, pour toute sous-variété singulière $\gamma : \mathscr{D} \to \mathscr{X}$ sera définie la sous-variété singulière

$$\gamma\,|_{\partial\mathscr{D}} : \partial\mathscr{D} \to \mathscr{X}.$$

Cette sous-variété singulière sera appelée *bord* de la sous-variété singulière γ et désignée par $\partial\gamma$.

Théorème 1″ (théorème de Stokes pour sous-variétés singulières). *Pour toute sous-variété singulière γ à n dimensions d'une variété $\mathscr{X}$ et toute forme ω à support borné de degré $n - 1$ sur γ on a la formule*

$$\int_{\gamma} d\omega = \int_{\gamma} \omega \quad \textit{sur } \mathscr{X}. \tag{17}$$

Démonstration. Il suffit d'appliquer le théorème 1′ à la forme $\gamma^*\omega$ sur $\mathscr{D}$. □

Le théorème 1‴ se ramène au théorème 1′ lorsque γ est l'application identique $\mathrm{id}: \mathscr{D} \to \mathscr{D}$, et au théorème 1″ lorsque γ est une injection de $\mathscr{D}$ dans $\mathscr{X}$.

* * *

Les intégrales étendues aux sous-variétés singulières à une dimension de l'espace $\mathbb{R}^n$, $n \geqslant 2$, admettent une interprétation élémentaire. Supposons pour fixer les idées que $n = 2$ et que $\mathscr{D}$ est un segment $[a, b]$. Donc, la variété singulière $\gamma: \mathscr{D} \to \mathbb{R}^2$ n'est autre que la courbe différentiable plane

$$x = x(t), \quad y = y(t), \quad a \leqslant t \leqslant b \tag{18}$$

(qui généralement n'est soumise à aucune condition de régularité). Un déchiffrage élémentaire des définitions montre que pour toute forme $P\,\mathrm{d}x + Q\,\mathrm{d}y$ sur $\mathbb{R}^2$ l'intégrale le long de la courbe (18) est définie par la formule

$$\int_\gamma P\,\mathrm{d}x + Q\,\mathrm{d}y = \int_a^b [P(x(t), y(t))\,x'(t) + Q(x(t), y(t))\,y'(t)]\,\mathrm{d}t, \tag{19}$$

formule qui peut servir de définition à cette intégrale. Les intégrales de cette nature interviennent fréquemment dans les problèmes d'analyse et s'appellent *intégrales curvilignes de seconde espèce*.

Pour $n = 3$ elles sont de la forme

$$\int_\gamma P\,\mathrm{d}x + Q\,\mathrm{d}y + R\,\mathrm{d}z = \int_a^b [P(x(t), y(t), z(t))\,x'(t) + \\ + Q(x(t), y(t), z(t))\,y'(t) + R(x(t), y(t), z(t))\,z'(t)]\,\mathrm{d}t,$$

idem pour tout n.

[Les intégrales curvilignes de première espèce ont été introduites dans la leçon précédente. Ce sont des intégrales de densités et non de formes.]

On remarquera que l'intégrale (19) a un sens aussi bien pour les courbes différentiables que pour les courbes différentiables par morceaux.

Lorsque $\mathscr{D}$ est un cercle orienté S^1, la variété singulière γ s'appelle *co rbe fermée*. En se donnant un segment $[a, b]$, un point $s_0 \in S^1$ et une application $\alpha: [a, b] \to S^1$ envoyant les points a et b en s_0 et appliquant difféomorphiquement (en préservant l'orientation) l'intervalle $]a, b[$ sur l'arc $S^1 \setminus \{s_0\}$, on peut identifier une telle courbe à une courbe $\beta = \gamma \circ \alpha: [a, b] \to \mathbb{R}^n$ telle que $\beta(a) = \beta(b)$.

(Pour cette raison, les courbes $\beta : [a. b] \to \mathbb{R}^n$ telles que $\beta(a) = \beta(b)$ sont dites aussi *courbes fermées*; comparer avec la leçon 1.)

Cette identification est compatible avec l'intégration, i.e. l'intégrale le long de toute courbe fermée γ (qui est généralement désignée par $\oint_\gamma$ pour exprimer que la courbe γ est fermée) est égale à l'intégrale curviligne le long de β.

Lorsque $\mathscr{D}$ est composée de plusieurs cercles orientés, la variété singulière γ est un système $\gamma_1, \ldots, \gamma_n$ de courbes fermées, et l'intégrale correspondante est justiciable de la formule

$$\oint_\gamma = \oint_{\gamma_1} + \quad + \oint_{\gamma_n} .$$

En particulier, pour tout domaine compact plan $\mathscr{D} \subset \mathbb{R}^2$ à frontière régulière (ou régulière par morceaux dans un sens notoire) et toute forme $P\,dx + Q\,dy$ sur $\mathbb{R}^2$, ceci identifie l'intégrale $\oint_{\partial\mathscr{D}} P\,dx + Q\,dy$ de la formule de Green (7) à l'intégrale curviligne $\oint_{\partial\mathscr{D}} P\,dx + Q\,dy$.

Donc, les deux intégrales de la formule de Green (7) admettent une signification élémentaire, ce qui n'est pas le cas, par exemple, de l'intégrale du second membre de la formule de Gauss-Ostrogradski (6) ou du premier membre de la formule de Stokes (14).

LEÇON 28

Opérateurs d'analyse vectorielle. — Corollaires de l'identité $d \circ d = 0$. — Corollaires de la formule de dérivation d'un produit. — Opérateurs de Laplace et de Beltrami. — Flux d'un champ de vecteurs. — Formule de Gauss-Ostrogradski pour la divergence et formules de Green.—La divergence en tant que densité de sources.— Formule de Stokes pour la circulation. — Formule de Gauss-Ostrogradski pour le rotationnel. — Formule généralisée de Gauss-Ostrogradski.

On étudie dans cette leçon un cas de formes sur l'espace $\mathbb{R}^3$, qui est particulièrement important pour les applications. Ce cas a en fait été partiellement étudié à la leçon I.22. Pour la commodité, nous répétons notre exposé en le complétant avec les corollaires de la formule de Stokes.

L'existence dans $\mathbb{R}^3$ d'un système de coordonnées fixe et d'une métrique euclidienne standard nous permet d'identifier les formes différentielles linéaires $P\,dx + Q\,dy + R\,dz$ sur $\mathbb{R}^3$ (ainsi que les formes $P\,dy \wedge dz + Q\,dz \wedge dx + R\,dx \wedge dy$ du second degré) aux champs de vecteurs

$$\mathbf{u} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}, \tag{1}$$

où $\mathbf{i}$, $\mathbf{j}$ et $\mathbf{k}$ sont les vecteurs de base de $\mathbb{R}^3$. Ceci nous conduit à une théorie assez féconde connue sous le nom d'a n a l y s e v e c t o r i e l l e (o u t h é o r i e d u c h a m p). Bien qu'elle soit de portée théorique minimale, l'analyse vectorielle sera exposée avec suffisamment de détails, car elle joue un rôle important dans les applications physico-techniques liées à la dynamique des fluides et à l'électromagnétisme (il suffit de signaler que les équations de Maxwell du champ électromagnétique sont décrites le plus élégamment — exception faite du formalisme quadridimensionnel de la théorie de relativité spéciale — par les opérateurs différentiels d'analyse vectorielle).

Le principal champ d'action de l'analyse vectorielle est l'algèbre des fonctions différentiables F définies sur un ensemble ouvert $W \subset \mathbb{R}^3$ (qui sera fixé dans la suite), et le F-module $\mathfrak{a}$ des champs de vecteurs sur W.

L'identification du module $\mathfrak{a}$ aux modules Ω^1 et Ω^2 des formes différentielles de degré 1 et 2 (et l'identification de l'algèbre F, traitée comme un module sur elle-même, au module Ω^3 des formes $f\,dx \wedge dy \wedge dz$ de degré 3) nous permet de traiter les opérateurs de dérivation extérieure

$$\mathsf{F} = \Omega^0 \xrightarrow{d} \Omega^1 \xrightarrow{d} \Omega^2 \xrightarrow{d} \Omega^3$$

comme des opérateurs

$$\mathsf{F} \to \alpha \to \alpha \to \mathsf{F}.$$

Le premier de ces opérateurs (de F dans α) est désigné par le symbole grad. Il associe à une fonction différentiable arbitraire F le champ de vecteurs

$$\text{(2)} \qquad \operatorname{grad} F = \frac{\partial F}{\partial x}\,\mathbf{i} + \frac{\partial F}{\partial y}\,\mathbf{j} + \frac{\partial F}{\partial z}\,\mathbf{k}$$

appelé *gradient* de la fonction F.

Le deuxième opérateur (de α dans α) est désigné par rot. Il envoie le champ de vecteurs (1) dans le champ

$$\text{(3)} \qquad \operatorname{rot} \mathbf{u} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}$$

appelé *rotationnel* du champ $\mathbf{u}$.

La formule (3) peut être mise sous la forme simple à retenir

$$\text{(4)} \qquad \operatorname{rot} \mathbf{u} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & R \end{vmatrix}.$$

Naguère, on écrivait curl $\mathbf{u}$ pour rot $\mathbf{u}$, mais cette notation est tombée en désuétude.

Le troisième opérateur (de α dans F) est désigné par div. Il transforme le champ de vecteurs (1) en la fonction

$$\text{(5)} \qquad \operatorname{div} \mathbf{u} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z},$$

appelée *divergence* du champ $\mathbf{u}$.

Outre ces opérateurs (et la multiplication par une fonction) sur α sont définis le produit scalaire et le produit vectoriel qui aux champs de vecteurs

$$\mathbf{u} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k} \quad \text{et} \quad \mathbf{v} = X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k}$$

associent respectivement la fonction

$$\mathbf{uv} = PX + QY + RZ$$

et le champ de vecteurs

$$\mathbf{u} \times \mathbf{v} = (QZ - RY)\,\mathbf{i} + (RX - PZ)\mathbf{j} + (PY - QX)\,\mathbf{k} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ P & Q & R \\ X & Y & Z \end{vmatrix}.$$

Ces opérateurs (et opérations) sont reliés entre eux par des multiples identités.

* * *

Tout d'abord l'identité $d \circ d = 0$ nous donne deux identités

$$\text{rot grad } F = 0 \quad \text{et} \quad \text{div rot } \mathbf{u} = 0, \tag{6}$$

qui du reste se vérifient aisément par un calcul direct.

Les champs de la forme grad F s'appellent *champs potentiels*, et ceux de la forme rot $\mathbf{u}$, *solénoïdaux*. Si $\mathbf{u} = \text{grad } F$, la fonction F s'appelle *potentiel* du champ $\mathbf{u}$, et si $\mathbf{v} = \text{rot } \mathbf{u}$, le champ $\mathbf{u}$ s'appelle *potentiel vecteur* du champ $\mathbf{v}$. Si W est connexe, le potentiel F est défini par le champ $\mathbf{u}$ à une constante additive près.

Un champ $\mathbf{u}$ est dit *irrotationnel* si rot $\mathbf{u} = 0$, et *sans sources* si div $\mathbf{u} = 0$. Le potentiel vecteur $\mathbf{u}$ d'un champ rot $\mathbf{u}$ est défini à un champ irrotationnel près.

En vertu de l'identité (6) *tout champ potentiel est irrotationnel, et tout champ solénoïdal est un champ sans sources.* [Ceci étant, l'espace quotient de l'espace vectoriel des champs irrotationnels par le sous-espace des champs potentiels n'est autre que le groupe de cohomologie de de Rham H^1W du domaine W, et l'espace quotient des champs sans sources par le sous-espace des champs solénoïdaux, le groupe H^2W. Donc, les champs irrotationnels sont potentiels si et seulement si $H^1W = 0$, et les champs sans sources sont solénoïdaux si et seulement si $H^2W = 0$.]

Exemple 1. Un champ de la forme

$$\mathbf{u} = f(r)\,\mathbf{r}, \quad \mathbf{r} \neq 0,$$

où $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ et $r = |\mathbf{r}| = \sqrt{x^2 + y^2 + z^2}$, s'appelle *champ central.* Pour ce champ,

$$P = f(r)\,x, \quad Q = f(r)\,y, \quad R = f(r)\,z$$

(et $W = \mathbb{R}^3 \setminus \{0\}$). Puisque

$$\frac{\partial r}{\partial x} = \frac{x}{r}, \quad \frac{\partial r}{\partial y} = \frac{y}{r}, \quad \frac{\partial r}{\partial z} = \frac{z}{r},$$

la matrice

$$\begin{Vmatrix} \dfrac{\partial P}{\partial x} & \dfrac{\partial P}{\partial y} & \dfrac{\partial P}{\partial z} \\ \dfrac{\partial Q}{\partial x} & \dfrac{\partial Q}{\partial y} & \dfrac{\partial Q}{\partial z} \\ \dfrac{\partial R}{\partial x} & \dfrac{\partial R}{\partial y} & \dfrac{\partial R}{\partial z} \end{Vmatrix}$$

est de la forme

$$(7)\qquad \begin{Vmatrix} f'(r)\frac{x^2}{r}+f(r) & f'(r)\frac{xy}{r} & f'(r)\frac{xz}{r} \\ f'(r)\frac{xy}{r} & f'(r)\frac{y^2}{r}+f(r) & f'(r)\frac{yz}{r} \\ f'(r)\frac{xz}{r} & f'(r)\frac{yz}{r} & f'(r)\frac{z^2}{r}+f(r) \end{Vmatrix}.$$

La symétrie de cette matrice exprime que rot $\mathbf{u} = 0$, i.e. *tout champ central est irrotationnel.* Bien plus, comme $H^1(\mathbb{R}^3 \setminus \{0\}) = H^1(\mathbf{S}^2) = 0$, *tout champ central est potentiel.* (Le potentiel correspondant F est défini par la formule

$$F(\mathbf{r}) = \int_1^r rf(r)\,\mathrm{d}r,$$

où de toute évidence la forme inférieure d'intégration peut être tout autre nombre strictement positif.)

Si en particulier $f(r) = -\frac{m_0}{r^3}$ et donc $|\mathbf{u}| = \frac{m_0}{r^2}$ (c h a m p g r a v i t a t i o n n e l d'u n p o i n t m a t é r i e l d e m a s s e m_0), alors $F(\mathbf{r}) = \frac{m_0}{r}$ (à une constante près). Ce potentiel est dit *newtonien.*

La divergence d'un champ central est égale à la somme

$$f'(r)\frac{x^2+y^2+z^2}{r}+3f(r) = rf'(r)+3f(r)$$

des éléments diagonaux de la matrice (7). La résolution de l'équation différentielle $rf'(r) + 3f(r) = 0$ nous donne $f(r) = \frac{C}{r^3}$. Donc, *un champ central est un champ gravitationnel (le champ d'un potentiel newtonien) si et seulement s'il est sans source* (dans le domaine $\mathbb{R}^3 \setminus \{0\}$ où il est défini).

* * *

Le produit extérieur $\omega \wedge \theta$ des formes différentielles ω et θ se transforme en vertu de nos identifications en le produit vectoriel des champs ω et θ si deg $\omega = 1$ et deg $\theta = 1$, en le produit scalaire de ces champs si deg $\omega = 1$ et deg $\theta = 2$ (et en le produit du champ θ par la fonction ω si deg $\omega = 0$). Donc, la formule

$$\mathrm{d}(f\omega) = f\,\mathrm{d}\omega + \mathrm{d}f \wedge \omega$$

nous donne les relations

$$\text{grad}\,(FG) = F\,\text{grad}\,G + G\,\text{grad}\,F, \tag{8}$$

$$\text{rot}\,(F\mathbf{u}) = F\,\text{rot}\,\mathbf{u} + \text{grad}\,F \times \mathbf{u}, \tag{9}$$

$$\text{div}\,(F\mathbf{u}) = F\,\text{div}\,\mathbf{u} + \text{grad}\,F \cdot \mathbf{u}, \tag{10}$$

et la formule

$$d\,(\omega \wedge \theta) = d\omega \wedge \theta - \omega \wedge d\omega, \qquad \omega, \theta \in \Omega^1,$$

la relation

$$\text{div}\,(\mathbf{u} \times \mathbf{v}) = (\text{rot}\,\mathbf{u})\,\mathbf{v} - \mathbf{u}\,\text{rot}\,\mathbf{v}. \tag{11}$$

Il est évident que les formules (8) à (11) s'obtiennent sans peine par un calcul direct.

La formule (4) nous suggère d'introduire le champ de vecteurs symbolique

$$\nabla = \frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}.$$

Le champ rot $\mathbf{u}$ sera alors le produit vectoriel $\nabla \times \mathbf{u}$ du champ ∇ par le champ $\mathbf{u}$:

$$\text{rot}\,\mathbf{u} = \nabla \times \mathbf{u},$$

la fonction div $\mathbf{u}$, le produit scalaire $\nabla\mathbf{u}$ du champ ∇ par le champ $\mathbf{u}$:

$$\text{div}\,\mathbf{u} = \nabla\mathbf{u},$$

et le champ grad F, si l'on écrit le facteur scalaire à droite, le produit ∇F de la fonction F par le champ ∇:

$$\text{grad}\,F = \nabla F.$$

Ces représentations des opérateurs rot, div et grad par l'opérateur symbolique ∇ permettent d'écrire les relations (8) à (11) sous la forme d une seule identité facile à retenir

$$\nabla \textcircled{1} (\alpha \textcircled{1} \gamma) = \nabla \textcircled{1} (\overset{\downarrow}{\alpha} \textcircled{2} \beta) + \nabla \textcircled{1} (\alpha \textcircled{2} \overset{\downarrow}{\beta}), \tag{12}$$

où α et β sont soit des fonctions, soit des champs, ① et ② deux des trois produits possibles (produit par une fonction, produit scalaire et produit vectoriel), la flèche $\downarrow$ indiquant le facteur soumis à l'action de l'opérateur ∇. En effet, pour $\alpha = F$ et $\beta = G$, on retrouve visiblement la formule (8), et pour $\alpha = F$ et $\beta = \mathbf{u}$, les formules (9) et (10) selon le produit choisi. Pour $\alpha = \mathbf{u}$ et $\beta = \mathbf{v}$, et sous réserve que ① soit le produit scalaire et ② le produit vectoriel, la formule (12) devient

$$\nabla (\mathbf{u} \times \mathbf{v}) = \nabla (\overset{\downarrow}{\mathbf{u}} \times \mathbf{v}) + \nabla (\mathbf{u} \times \overset{\downarrow}{\mathbf{v}}).$$

Mais en vertu de l'antisymétrie du produit mixte (et la commutativité du produit scalaire)

$$\nabla(\overset{\downarrow}{\mathbf{u}} \times \mathbf{v}) = \nabla \overset{\downarrow}{\mathbf{u}}\mathbf{v} = \mathbf{v}\nabla\overset{\downarrow}{\mathbf{u}} = \mathbf{v}\nabla\mathbf{u} = \mathbf{v}(\nabla \times \mathbf{u}) = (\nabla \times \mathbf{u})\,\mathbf{v}$$

et

$$\nabla(\mathbf{u} \times \overset{\downarrow}{\mathbf{v}}) = \nabla\mathbf{u}\overset{\downarrow}{\mathbf{v}} = -\mathbf{u}\nabla\overset{\downarrow}{\mathbf{v}} = -\mathbf{u}\nabla\mathbf{v} = -\mathbf{u}(\nabla \times \mathbf{v}).$$

Donc, dans ce cas la formule (12) se ramène à la formule (11).

Remarquons que pour $\alpha = \mathbf{u}$ et $\beta = \mathbf{v}$ il reste à traiter deux cas dans la formule (12):

$$\nabla(\mathbf{uv}) = \nabla(\overset{\downarrow}{\mathbf{u}}\mathbf{v}) + \nabla(\mathbf{u}\overset{\downarrow}{\mathbf{v}}) \tag{13}$$

et

$$\nabla \times (\mathbf{u} \times \mathbf{v}) = \nabla \times (\overset{\downarrow}{\mathbf{u}} \times \mathbf{v}) + \nabla \times (\mathbf{u} \times \overset{\downarrow}{\mathbf{v}}). \tag{14}$$

Pour déchiffrer ces formules, on conviendra pour tout champ de vecteurs $\mathbf{a} = A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$ de comprendre par $\mathbf{a}\nabla$ l'opérateur $\alpha \to \mathsf{F}$ agissant sur le champ $\mathbf{u} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ à l'aide de la formule

$$(\mathbf{a}\nabla)\,\mathbf{u} = A\frac{\partial P}{\partial x} + B\frac{\partial Q}{\partial y} + C\frac{\partial R}{\partial z}$$

(ce faisant, on renonce à la commutativité du produit scalaire lorsqu'il s'agit du vecteur symbolique ∇). En développant les doubles produits vectoriels $\nabla \times (\overset{\downarrow}{\mathbf{u}} \times \mathbf{v})$ et $\nabla \times (\mathbf{u} \times \overset{\downarrow}{\mathbf{v}})$ (à l'aide de la formule classique $\mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = (\mathbf{cb})\,\mathbf{a} - (\mathbf{ac})\,\mathbf{b}$; cf. formule (19) de la leçon I.22), on trouve que

$$\nabla \times (\overset{\downarrow}{\mathbf{u}} \times \mathbf{v}) = (\mathbf{v}\nabla)\,\mathbf{u} - (\nabla\mathbf{u})\,\mathbf{v} = (\mathbf{v}\nabla)\,\mathbf{u} - (\operatorname{div}\mathbf{u})\,\mathbf{v}$$

et

$$\nabla \times (\mathbf{u} \times \overset{\downarrow}{\mathbf{v}}) = (\nabla\mathbf{v})\,\mathbf{u} - (\mathbf{u}\nabla)\,\mathbf{v} = (\operatorname{div}\mathbf{v})\,\mathbf{u} - (\mathbf{u}\nabla)\,\mathbf{v}.$$

Puisque par définition $\nabla \times (\mathbf{u} \times \mathbf{v}) = \operatorname{rot}(\mathbf{u} \times \mathbf{v})$, la formule (14) devient

$$\operatorname{rot}(\mathbf{u} \times \mathbf{v}) = (\mathbf{v}\nabla)\mathbf{u} - (\mathbf{u}\nabla)\,\mathbf{v} + (\operatorname{div}\mathbf{v})\mathbf{u} - (\operatorname{div}\mathbf{u})\,\mathbf{v}. \tag{15}$$

De façon analogue, en appliquant aux champs $\mathbf{u} \times \operatorname{rot}\mathbf{v} = \mathbf{u} \times (\nabla \times \mathbf{v})$ et $\mathbf{v} \times \operatorname{rot}\mathbf{u} = \mathbf{v} \times (\nabla \times \mathbf{u})$ la formule du double produit vectoriel mise sous la forme $\mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a}\,(\mathbf{cb}) - (\mathbf{ca})\,\mathbf{b}$, on obtient les relations

$$\mathbf{u} \times \operatorname{rot}\mathbf{v} = \nabla(\mathbf{u}\overset{\downarrow}{\mathbf{v}}) - (\mathbf{u}\nabla)\,\mathbf{v}$$

et

$$\mathbf{v} \times \operatorname{rot} \mathbf{u} = \nabla (\overset{\downarrow}{\mathbf{u}}\mathbf{v}) - (\mathbf{v}\nabla)\, \mathbf{u},$$

grâce auxquelles la formule (13) devient

(16) $\operatorname{grad} (\mathbf{uv}) = \mathbf{u} \times \operatorname{rot} \mathbf{u} + \mathbf{v} \operatorname{rot} \mathbf{u} + (\mathbf{v}\nabla)\, \mathbf{u} + (\mathbf{u}\nabla)\, \mathbf{v}.$

Il est remarquable, comme l'atteste un calcul direct, que les formules (15) et (16) sont effectivement valables (de sorte que la formule (12) est toujours vraie quand elle a un sens).

[La complexité des formules (15) et (16) s'explique par le fait que les valeurs de l'opérateur rot sur le champ de vecteurs $\mathbf{u} \times \mathbf{v}$ et de l'opérateur grad sur la fonction $\mathbf{uv}$ n'admettent pas d'interprétation directe en termes de dérivation extérieure des formes.]

Exemple 2. Soit $\mathbf{a}$ un vecteur constant. Calculons le champ $\operatorname{grad}(\mathbf{ar})$, où comme toujours $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Il est clair que $\operatorname{rot} \mathbf{a} = 0$ et $(\mathbf{r}\nabla)\mathbf{a} = 0$. De plus, un calcul direct montre que $\operatorname{rot} \mathbf{r} = 0$ (le champ $\mathbf{r}$ est central) et $(\mathbf{a}\nabla)\mathbf{r} = \mathbf{a}$. Donc, $\operatorname{grad}(\mathbf{ar}) = \mathbf{a}$.

* * *

En combinant les opérateurs grad, rot et div, on obtient les opérateurs

$$\operatorname{rot} \circ \operatorname{grad} : \mathsf{F} \to \mathfrak{a}, \quad \operatorname{div} \circ \operatorname{grad} : \mathsf{F} \to \mathsf{F},$$

$$\operatorname{grad} \circ \operatorname{div} : \mathfrak{a} \to \mathfrak{a}, \quad \operatorname{rot} \circ \operatorname{rot} : \mathfrak{a} \to \mathfrak{a},$$

$$\operatorname{div} \circ \operatorname{rot} : \mathfrak{a} \to \mathsf{F}.$$

On sait déjà que les opérateurs $\operatorname{rot} \circ \operatorname{grad}$ et $\operatorname{div} \circ \operatorname{rot}$ sont identiquement nuls. [On remarquera que l'opérateur ∇ ramène cette affirmation à celle que le produit vectoriel de deux vecteurs identiques est nul: $(\operatorname{rot} \circ \operatorname{grad})\ F = \nabla \times \nabla F = (\nabla \times \nabla)\, F = 0$ et $(\operatorname{div} \circ \operatorname{rot})\, \mathbf{u} = \nabla (\nabla \times \mathbf{u}) = (\nabla \times \nabla)\mathbf{u} = 0$.]

Le plus intéressant des autres opérateurs est

$$\operatorname{div} \circ \operatorname{grad} = \nabla^2.$$

Cet opérateur est désigné par Δ et porte le nom *d'opérateur de Laplace* (ou *laplacien*). Il associe à toute fonction F la fonction

$$\Delta F = \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} + \frac{\partial^2 F}{\partial z^2}$$

et permet d'écrire les plus importantes équations de physique mathématique.

On peut appliquer l'opérateur Δ aux champs de vecteurs en le faisant agir sur chaque composante séparément: si $\mathbf{u} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$, on a par définition

$$\Delta \mathbf{u} = (\Delta P)\, \mathbf{i} + (\Delta Q)\, \mathbf{j} + (\Delta R)\, \mathbf{k}.$$

Alors

$$\text{rot} \circ \text{rot} = \text{grad} \circ \text{div} - \Delta.$$

En effet, d'après la formule du double produit vectoriel,

$$\text{rot}\,(\text{rot}\,\mathbf{u}) = \nabla \times (\nabla \times \mathbf{u}) = \nabla\,(\nabla \mathbf{u}) - (\nabla\nabla)\mathbf{u}\,. \square$$

[Ceci n'est certes pas une démonstration, mais une simple confirmation euristique. La véritable démonstration qui est laissée au soin du lecteur n'est donnée que par un calcul direct à l'aide des formules (2), (3) et (5).]

On pourrait composer d'autres expressions différentielles. Par exemple, pour deux fonctions F et G quelconques, est défini le produit scalaire de leurs gradients:

$$\text{grad}\,F\,\text{grad}\,G = \frac{\partial F}{\partial x}\frac{\partial G}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial G}{\partial y} + \frac{\partial F}{\partial z}\frac{\partial G}{\partial z}.$$

On le désigne par le symbole $\Delta\,(F, G)$ et on l'appelle *paramètre différentiel mixte de Beltrami des fonctions F et G*. En particulier, pour $F = G$, on obtient le carré scalaire du gradient:

$$\Delta\,(F, F) = (\text{grad}\,F)^2 = \left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2 + \left(\frac{\partial F}{\partial z}\right)^2.$$

On le désigne par le symbole $\Delta_1 F$ et on l'appelle *premier paramètre différentiel de Beltrami de la fonction F* (cf. leçon 3).

Le produit mixte des gradients de trois fonctions s'appelle *paramètre différentiel de Darboux*. Ce terme n'est du reste pratiquement pas utilisé aujourd'hui, car ce produit mixte n'est autre que le jacobien de l'application définie par les trois fonctions données.

* * *

Les opérateurs d'analyse vectorielle nous permettent de représenter la formule de Gauss-Ostrogradski (et la formule de Stokes) sous une forme commode et concise.

Soit $\mathscr{D}$ une surface orientée à deux dimensions dans $\mathbb{R}^3$ (avec ou sans bord). On sait (cf. remarque 2 de la leçon 27) que l'orientation de la surface $\mathscr{D}$ est définie par un champ différentiable de vecteurs normaux non nuls $\mathbf{n} = \mathbf{n}\,(M)$, qui peut être normé.

Un champ de vecteurs $\mathbf{a}$ défini sur un ensemble ouvert W contenant une surface $\mathscr{D}$ définit sur $\mathscr{D}$ une fonction a_n associant à tout point $M \in \mathscr{D}$ la projection $a_n\,(M) = \mathbf{a}\,(M)\,\mathbf{n}\,(M)$ du vecteur $\mathbf{a}\,(M)$ sur la direction du vecteur $\mathbf{n}\,(M)$ (qui est appelée *composante normale* du vecteur $\mathbf{a}\,(M)$).

Définition 1. L intégrale

$$\int_{\mathscr{D}} a_n\,d\sigma \tag{17}$$

s'appelle *flux* du champ **a** à travers la surface $\mathscr{D}$. (Ici $d\sigma$ est l'élément d'aire de $\mathscr{D}$; cf. exemple 2 de la leçon 24.)

Signalons que l'intégrale de la formule (17) est une intégrale de première espèce (de la densité).

Si le champ **a** est le champ de vitesses d'un fluide, le flux (17) est égal à la quantité de fluide traversant la surface $\mathscr{D}$ en une unité de temps.

Si la surface $\mathscr{D}$ est élémentaire, pour toute paramétrisation $\mathbf{r} = \mathbf{r}(u, v)$ *compatible avec l'orientation*, i.e. telle que les vecteurs $\mathbf{r}_u$ et $\mathbf{r}_v$ forment une base positivement orientée sur le plan tangent, le vecteur normal **n** est donné par la formule

$$\mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{\sqrt{EG - F^2}}$$

(comparer à la leçon 4). D'autre part, de la leçon 3 on sait que l'élément d'aire d'une surface est donné par la formule $d\sigma = \sqrt{EG - F^2}\,du\,dv$. Donc, le flux (17) est justiciable de la formule

$$(18) \qquad \int_{\mathscr{D}} a_n \, d\sigma = \iint_U (\mathbf{a}\mathbf{r}_u\mathbf{r}_v)\, du\, dv = \iint_U \begin{Vmatrix} P & Q & R \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{Vmatrix} du\, dv,$$

où P, Q et R sont les coordonnées du vecteur **a**, et U le domaine du plan $\mathbb{R}^2$ sur lequel est définie la paramétrisation $\mathbf{r} = \mathbf{r}(u, v)$. Pour calculer le flux à travers une surface non élémentaire, il faut la partager en surfaces élémentaires et appliquer la formule (18) à chacune d'elles.

Rappelons maintenant qu'au champ $\mathbf{a} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ est associée la forme différentielle

$$P\, dy \wedge dz + Q\, dz \wedge dx + R\, dx \wedge dy,$$

et à cette forme, l'intégrale

$$(19) \qquad \int_{\mathscr{D}} P\, dy \wedge dz + Q\, dz \wedge dx + R\, dx \wedge dy =$$

$$= \iint_{\mathscr{D}} P\, dy\, dz + Q\, dz\, dx + R\, dx\, dy$$

(cf. formule (13) de la leçon 27). Lorsque la surface $\mathscr{D}$ est élémentaire et paramétrée (par une paramétrisation $\mathbf{r} = \mathbf{r}(u, v)$ définie sur un ouvert $U \subset \mathbb{R}^2$), en déchiffrant la définition de l'intégrale (19),

on trouve immédiatement qu'elle s'exprime par la formule

$$\iint_U \left[P\begin{vmatrix} y_u & z_u \\ y_v & z_v \end{vmatrix} - Q\begin{vmatrix} x_u & z_u \\ x_v & z_v \end{vmatrix} + R\begin{vmatrix} x_u & y_u \\ x_v & y_v \end{vmatrix}\right] du\, dv =$$

$$= \iint_U \begin{vmatrix} P & Q & R \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix} du\, dv.$$

En comparant cette formule à la formule (18), on constate que *le flux* (17) *du champ* **a** *à travers la surface* $\mathscr{D}$ *s'exprime à l'aide de l'intégrale* (19):

$$\int_{\mathscr{D}} a_n\, d\sigma = \iint_{\mathscr{D}} P\, dy\, dz + Q\, dz\, dx + R\, dx\, dy. \tag{20}$$

[Cette formule qui a été prouvée pour les surfaces élémentaires est valable par additivité pour toutes surfaces.]

On remarquera que la formule (20) exprime le flux à l'aide d'une intégrale de surface de d e u x i è m e e s p è c e.

* * *

Lorsque la surface est le bord ∂D d'un domaine à frontière régulière, l'intégrale (19) n'est autre que l'intégrale du second membre de la formule de Gauss-Ostrogradski (cf. formule (6) de la leçon 27). L'intégrant du premier membre de cette formule n'étant autre que div **a**, on voit que la formule de Gauss-Ostrogradski peut être écrite sous la forme suivante:

$$\iiint_D \operatorname{div} \mathbf{a}\, dV = \oiint_{\partial D} a_n\, d\sigma, \tag{21}$$

où dV désigne $dx\, dy\, dz$ (et le signe $\oiint$ exprime que l'intégration est étendue à une surface fermée). Textuellement: *le flux d'un champ de vecteurs à travers la frontière d'un domaine est égal à l'intégrale de la divergence de ce champ étendue à ce domaine.*

Lorsque $\mathbf{a} = \operatorname{grad} F$, la composante normale $a_n = \operatorname{grad} F \cdot \mathbf{n}$ est désignée par $\frac{\partial F}{\partial n}$ et s'appelle *dérivée de la fonction F suivant la normale* (ou plus brièvement *dérivée normale*). Comme $\operatorname{div} \mathbf{a} = \Delta F$, la formule (21) devient alors

$$\iiint_D \Delta F\, dV = \oiint_{\partial D} \frac{\partial F}{\partial n}\, d\sigma. \tag{22}$$

Pour $\mathbf{a} = F \operatorname{grad} G$, on a en vertu de la formule (10)

$$\operatorname{div} \mathbf{a} = F \operatorname{div} \operatorname{grad} G + \operatorname{grad} F \cdot \operatorname{grad} G = F \cdot \Delta G + \Delta (F, G)$$

et

$$a_n = F \operatorname{grad} G \cdot \mathbf{n} = F \frac{\partial G}{\partial n}.$$

Donc

$$\iiint_D [F \cdot \Delta G + \Delta (F, G)] \, dV = \oiint_{\partial D} F \frac{\partial G}{\partial n} \, d\sigma. \tag{23}$$

Cette formule s'appelle *première formule de Green*. Pour $F = G$ elle se transforme en la formule

$$\iiint_D (F \cdot \Delta F + \Delta_1 F) \, dV = \oiint_{\partial D} F \frac{\partial F}{\partial n} \, d\sigma,$$

et pour $F = 1$ en la formule (22).

En permutant les fonctions F et G dans la formule (23) et en retranchant la formule obtenue de la formule initiale, on est conduit à la formule

$$\iiint_D (F \cdot \Delta G - G \cdot \Delta F) \, dV = \oiint_{\partial D} \left(F \frac{\partial G}{\partial n} - G \frac{\partial F}{\partial n} \right) d\sigma, \tag{24}$$

connue sous le nom de *deuxième formule de Green*.

* * *

Pour le champ de vitesses d'un fluide, l'intégrale de droite de la formule (21) est égale à la quantité de fluide sortant du domaine D. Si cette quantité est positive, c'est que le domaine D présente des *sources*, si elle est négative, des *puits*. (Du reste il est commode de traiter les puits comme des sources à débit négatif.) L'intégrale exprime le débit total des sources du champ dans le domaine D. C'est pourquoi, en le divisant par le volume de D, on obtient le débit moyen de ces sources dans D.

Fixons un point M_0 dans le domaine D et considérons une suite $\{D_n\}$ de domaines contractiles en M_0, i.e. tels que $M_0 \in D_n$ et diam $D_n \to 0$ lorsque $n \to \infty$, où diam D_n est le diamètre du domaine D_n. La limite (si elle existe)

$$\lim_{n \to \infty} \frac{\iint_{\partial D_n} a_n \, d\sigma}{V_n}.$$

où V_n est le volume de D_n, n'est autre que la *densité des sources* du champ en M_0. Or en vertu de la formule (21), cette limite est éga-

le à

$$\lim_{n\to\infty} \frac{\iiint\limits_{D_n} \operatorname{div} \mathbf{a}\, dV}{V_n},$$

donc est égale à la valeur de div **a** en M_0. Par conséquent, *la divergence d'un champ de vecteurs n'est autre que la densité de ses sources*:

$$\operatorname{div} \mathbf{a} = \lim_{n\to\infty} \frac{\iint\limits_{D_n} a_n\, d\sigma}{V_n}.$$

Ceci nous donne une nouvelle définition de la divergence (et en particulier si $\mathbf{a} = \operatorname{grad} F$ une nouvelle définition de l'opérateur de Laplace ΔF) qui a l'avantage d'être suggestive physiquement (et qui explique pourquoi un champ tel que $\operatorname{div} \mathbf{a} = 0$ s'appelle champ sans sources).

Remarque 1. Il est utile de remarquer que cette définition de la divergence et de l'opérateur de Laplace est valable aussi pour les champs et fonctions non différentiables (voire même discontinues). Il suffit seulement que la limite correspondante existe.

Nous savons que (cf. exemple 1 ci-dessus) le champ

$$\mathbf{a} = -\frac{\mathbf{r}}{r^3} m_0 \tag{25}$$

d'une masse attractive ne possède pas de sources (dans le domaine $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ sur lequel il est défini). D'autre part, si S_ε est une sphère de rayon ε centrée en $\mathbf{0}$, alors

$$\oiint\limits_{S_\varepsilon} a_n\, d\sigma = -m_0 \oiint\limits_{S_\varepsilon} \frac{d\sigma}{r^2} = -\frac{m_0}{\varepsilon^2} \oiint\limits_{S_\varepsilon} d\sigma = -4\pi m_0, \tag{26}$$

puisque l'intégrale $\oiint\limits_{S_\varepsilon} d\sigma$ est égale à l'aire $4\pi\varepsilon^2$ de la sphère S_ε. Ceci exprime que *la source du champ gravitationnel* (25) *est située en* $\mathbf{0}$ *et son débit est proportionnel à la masse.* (Le signe moins de la formule (26) exprime qu'il s'agit en fait d'un puits.)

On peut donc dire que les masses attractives sont les sources (puits) des champs gravitationnels.

Remarque 2. En introduisant les potentiels newtoniens d'un corps pesant, on démontre que cette déduction est valable pour les champs gravitationnels de toute configuration, mais ceci déborde le cadre de notre exposé. [Ces problèmes sont étudiés en théorie du potentiel qui est exposée dans tout cours d'équations de physique mathématique.]

* * *

L'intégrale de surface figurant dans la formule de Stokes (cf. formule (14) de la leçon 27) n'est autre en vertu de la formule générale (20) que le flux

$$\iint_{\mathscr{D}} (\operatorname{rot} \mathbf{a})_n \, d\sigma$$

de rot **a** à travers la surface orientée $\mathscr{D}$. Pour donner une interprétation analogue à l'intégrale étendue au bord de la surface, on remarquera tout d'abord que la forme différentielle $\omega = P\,dx + Q\,dy + R\,dz$ associée au champ de vecteurs $\mathbf{a} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ peut être représentée sous forme du produit scalaire $\mathbf{a}\,d\mathbf{r}$, où

$$d\mathbf{r} = dx \cdot \mathbf{i} + dy \cdot \mathbf{j} + dz \cdot \mathbf{k},$$

et par conséquent, l'intégrale de la forme ω le long d'une courbe arbitraire γ peut être mise sous la forme

$$\int_{\gamma} \mathbf{a}\,d\mathbf{r}. \tag{27}$$

Supposons que la courbe γ est fermée.

Définition 2. L'intégrale (27) le long d'une courbe fermée γ s'appelle *circulation* du champ de vecteurs **a** le long de γ.

Pour souligner que la courbe γ est fermée, on désigne l'intégrale (27) par le symbole

$$\oint_{\gamma} \mathbf{a}\,d\mathbf{r}. \tag{28}$$

En particulier, la circulation (28) est définie le long du bord $\partial\mathscr{D}$ d'une surface orientable arbitraire $\mathscr{D}$ et n'est autre que l'intégrale curviligne figurant dans la formule de Stokes.

Nous voyons donc que la formule de Stokes pour les intégrales de surface dans $\mathbb{R}^3$ peut être mise sous la forme suivante:

$$\iint_{\mathscr{D}} (\operatorname{rot} \mathbf{a})_n \, d\sigma = \oint_{\partial\mathscr{D}} \mathbf{a}\,d\mathbf{r}. \tag{29}$$

Textuellement: *la circulation d'un champ de vecteurs le long du bord d'une surface est égale au flux du rotationnel de ce champ à travers cette surface.*

* * *

Pour rot **a** il existe une autre formule intégrale qui peut être traitée comme une variante de la formule de Gauss-Ostrogradski.

Supposons comme précédemment que $\mathscr{D}$ est une surface orientée dans $\mathbb{R}^3$ par le champ de vecteurs unitaires normaux **n** défini sur

elle. Pour tout champ de vecteurs $\mathbf{u} = X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k}$ sur $\mathscr{D}$ définissons une *intégrale de* $\mathbf{u}$ *sur* $\mathscr{D}$ par la formule

$$\iint_{\mathscr{D}} \mathbf{u}\, d\sigma = \left(\iint_{\mathscr{D}} X\, d\sigma\right)\mathbf{i} + \left(\iint_{\mathscr{D}} Y\, d\sigma\right)\mathbf{j} + \iint_{\mathscr{D}} Z\, d\sigma\Big)\mathbf{k}.$$

Cette définition est valable en particulier pour le champ $\mathbf{u} = \mathbf{n} \times \mathbf{a}$, où $\mathbf{a}$ est un champ de vecteurs arbitraire défini sur un domaine contenant la surface $\mathscr{D}$.

Définition 3. Lorsque la surface $\mathscr{D}$ est fermée (compacte et sans bord), l'intégrale

$$\iint_{\mathscr{D}} (\mathbf{n} \times \mathbf{a})\, d\sigma \tag{30}$$

s'appelle *circulation* du champ de vecteurs $\mathbf{a}$ étendue à la surface $\mathscr{D}$ et se note

$$\oiint_{\mathscr{D}} (\mathbf{n} \times \mathbf{a})\, d\sigma.$$

Si la surface $\mathscr{D}$ est élémentaire et $\mathbf{r} = \mathbf{r}(u, v)$ est une paramétrisation de $\mathscr{D}$, pour des raisons déjà connues on a

$$\iint_{\mathscr{D}} (\mathbf{n} \times \mathbf{a})\, d\sigma = \iint_{U} ((\mathbf{r}_u \times \mathbf{r}_v) \times \mathbf{a})\, du\, dv, \tag{31}$$

où U est un ensemble ouvert du plan $\mathbb{R}^2$ sur lequel est définie la paramétrisation $\mathbf{r} = \mathbf{r}(u, v)$.

Or

$$\begin{aligned}
(\mathbf{r}_u \times \mathbf{r}_v) \times \mathbf{a} &= (\mathbf{a}\mathbf{r}_u)\,\mathbf{r}_v - (\mathbf{a}\mathbf{r}_v)\,\mathbf{r}_u = \\
&= [(Px_u + Qy_u + Rz_u)\,x_v - (Px_v + Qy_v + Rz_v)\,x_u]\,\mathbf{i} + \\
&\quad + [(Px_u + Qy_u + Rz_u)\,y_v - (Px_v + Qy_v + Rz_v)\,y_u]\,\mathbf{j} + \\
&\quad + [(Px_u + Qy_u + Rz_u)\,z_v - (Px_v + Qy_v + Rz_v)\,z_u]\,\mathbf{k} = \\
&= \left(R\begin{vmatrix} z_u & z_v \\ x_u & x_v \end{vmatrix} - Q\begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}\right)\mathbf{i} + \left(P\begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} - R\begin{vmatrix} y_u & y_v \\ z_u & z_v \end{vmatrix}\right)\mathbf{j} + \\
&\quad + \left(Q\begin{vmatrix} y_u & y_v \\ z_u & z_v \end{vmatrix} - P\begin{vmatrix} z_u & z_v \\ x_u & x_v \end{vmatrix}\right)\mathbf{k}
\end{aligned}$$

et

$$\iint_{\mathscr{D}} R\,\mathrm{d}z\,\mathrm{d}x - Q\,\mathrm{d}x\,\mathrm{d}y = \iint_U \left(R \begin{vmatrix} z_u & z_v \\ x_u & x_v \end{vmatrix} - Q \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} \right) \mathrm{d}u\,\mathrm{d}v,$$

$$\iint_{\mathscr{D}} P\,\mathrm{d}x\,\mathrm{d}y - R\,\mathrm{d}y\,\mathrm{d}z = \iint_U \left(P \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} - R \begin{vmatrix} y_u & y_v \\ z_u & z_v \end{vmatrix} \right) \mathrm{d}u\,\mathrm{d}v,$$

$$\iint_{\mathscr{D}} Q\,\mathrm{d}y\,\mathrm{d}z - P\,\mathrm{d}z\,\mathrm{d}x = \iint_U \left(Q \begin{vmatrix} y_u & y_v \\ z_u & z_v \end{vmatrix} - P \begin{vmatrix} x_u & x_v \\ z_u & z_v \end{vmatrix} \right) \mathrm{d}u\,\mathrm{d}v.$$

Donc

$$\iint_{\mathscr{D}} (\mathbf{n} \times \mathbf{a})\,\mathrm{d}\sigma = \left(\iint_{\mathscr{D}} R\,\mathrm{d}z\,\mathrm{d}x - Q\,\mathrm{d}x\,\mathrm{d}y \right) \mathbf{i} +$$

$$+ \left(\iint_{\mathscr{D}} P\,\mathrm{d}x\,\mathrm{d}y - R\,\mathrm{d}y\,\mathrm{d}z \right) \mathbf{j} + \left(\iint_{\mathscr{D}} Q\,\mathrm{d}y\,\mathrm{d}z - P\,\mathrm{d}z\,\mathrm{d}x \right) \mathbf{k}.$$

Cette formule est valable par additivité pour une surface (non élémentaire) arbitraire $\mathscr{D}$.

Mais si $\mathscr{D} = \partial D$, où D est un domaine à frontière régulière, la formule de Gauss-Ostrogradski nous donne

$$\oiint_{\partial D} R\,\mathrm{d}z\,\mathrm{d}x - Q\,\mathrm{d}x\,\mathrm{d}y = \iiint_D \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \mathrm{d}V,$$

$$\oiint_{\partial D} P\,\mathrm{d}x\,\mathrm{d}y - R\,\mathrm{d}y\,\mathrm{d}z = \iiint_D \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) \mathrm{d}V,$$

$$\oiint_{\partial D} Q\,\mathrm{d}y\,\mathrm{d}z - P\,\mathrm{d}z\,\mathrm{d}x = \iiint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}V.$$

Comme

$$\operatorname{rot} \mathbf{a} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) \mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathbf{k}.$$

ceci prouve que *la circulation d'un champ de vecteurs le long de la frontière ∂D d'un domaine D est égale à l'intégrale du rotationnel de ce champ étendue à D*:

$$\oiint_{\partial D} (\mathbf{n} \times \mathbf{a})\,\mathrm{d}\sigma = \iiint_D \operatorname{rot} \mathbf{a}\,\mathrm{d}V. \tag{32}$$

[Il s'ensuit en particulier que le rotationnel d'un champ de vecteurs peut être traité comme la *densité de la circulation* de ce champ (ceci nous donne en particulier une nouvelle définition du rotationnel qui est valable pour les champs non différentiables).]

* * *

La formule (32) est un cas particulier d'une formule générale dont la démonstration est, comme il arrive souvent, bien plus courte que celle de son cas particulier (32).

Soit φ une application linéaire de $\mathbb{R}^3$ soit dans l'espace des fonctions F, soit dans l'espace des champs de vecteurs α. (La linéarité exprime que pour tout vecteur $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ de l'espace $\mathbb{R}^3$ on a l'égalité $\varphi(\mathbf{r}) = x\varphi(\mathbf{i}) + y\varphi(\mathbf{j}) + z\varphi(\mathbf{k})$. Associons à l'application φ une fonction (ou un champ) $\varphi(\nabla)$ définie par la formule

$$\varphi(\nabla) = \frac{\partial}{\partial x}\varphi(\mathbf{i}) + \frac{\partial}{\partial y}\varphi(\mathbf{j}) + \frac{\partial}{\partial z}\varphi(\mathbf{k}).$$

Si par exemple $\varphi(\mathbf{r}) = \mathbf{r}F$ (le produit du vecteur $\mathbf{r}$ par une fonction F), on a

$$\varphi(\nabla) = \nabla F = \frac{\partial F}{\partial x}\mathbf{i} + \frac{\partial F}{\partial y}\mathbf{j} + \frac{\partial F}{\partial z}\mathbf{k} = \operatorname{grad} F,$$

et si $\varphi(\mathbf{r}) = \mathbf{ra}$ ou $\varphi(\mathbf{r}) = \mathbf{r} \times \mathbf{a}$, où $\mathbf{a}$ est un champ, on a respectivement

$$\varphi(\nabla) = \nabla\mathbf{a} = \operatorname{div}\mathbf{a} \quad \text{et} \quad \varphi(\nabla) = \nabla \times \mathbf{a} = \operatorname{rot}\mathbf{a}.$$

Supposons pour fixer les idées que $\varphi: \mathbb{R}^3 \to \mathsf{F}$. Alors $\varphi(\nabla) = \operatorname{div}\mathbf{a}$, où $\mathbf{a} = \varphi(\mathbf{i})\,\mathbf{i} + \varphi(\mathbf{j})\,\mathbf{j} + \varphi(\mathbf{k})\,\mathbf{k}$, et donc en vertu de la formule de Gauss-Ostrogradski (20), pour tout domaine D on aura

$$\iiint_D \varphi(\nabla)\,\mathrm{d}V = \oiint_{\partial D} a_n\,\mathrm{d}\sigma.$$

Mais si $\mathbf{n} = \cos\alpha\cdot\mathbf{i} + \cos\beta\cdot\mathbf{j} + \cos\gamma\cdot\mathbf{k}$, alors

$$a_n = \mathbf{an} = \varphi(\mathbf{i})\cos\alpha + \varphi(\mathbf{j})\cos\beta + \varphi(\mathbf{k})\cos\gamma = \\ = \varphi(\cos\alpha\cdot\mathbf{i} + \cos\beta\cdot\mathbf{j} + \cos\gamma\cdot\mathbf{k}) = \varphi(\mathbf{n})$$

et donc

$$\iiint_D \varphi(\nabla)\,\mathrm{d}V = \oiint_{\partial D} \varphi(\mathbf{n})\,\mathrm{d}\sigma. \tag{33}$$

Si $\varphi: \mathbb{R}^3 \to \alpha$, en appliquant la formule (33) à chaque composante de l'application φ, on trouve immédiatement que cette formule est valable dans ce cas aussi.

La formule (33) s'appelle *formule généralisée de Gauss-Ostrogradski.*

Pour $\varphi(\mathbf{r}) = \mathbf{ra}$ elle se transforme en la formule (20), et pour $\varphi(\mathbf{r}) = \mathbf{r} \times \mathbf{a}$, en la formule (32).

LEÇON 29

Périodes des formes différentielles. — Simplexes, chaînes, cycles et et frontières singuliers. — Théorème de Stokes pour intégrales le long de chaînes. — Groupes d'homologie singulière. — Théorème de de Rham. — Groupes de cohomologie d'un complexe de chaînes. — Groupes de cohomologie singulière.

L'application du théorème de Stokes au calcul des groupes de cohomologie $H^m\mathscr{X}$ d'une variété séparée différentiable $\mathscr{X}$ (pas forcément sans bord) est basée sur le fait que pour toute forme différentielle fermée $\omega \in \Omega^m\mathscr{X}$ et toute sous-variété $\mathscr{Y} \subset \mathscr{X}$ orientée fermée (i.e. compacte et sans bord) à m dimensions, l'intégrale

$$I_{\mathscr{Y}}[\omega] = \int_{\mathscr{Y}} \omega \tag{1}$$

dépend uniquement de la classe de cohomologie $[\omega] \in H^m\mathscr{X}$ de cette forme (puisque, si ω est exacte, cette intégrale est nulle d'après la formule (4) de la leçon 27). Donc, la formule (1) définit de façon unique une application (visiblement linéaire)

$$I_{\mathscr{Y}} : H^m\mathscr{X} \to \mathbb{R}, \quad [\omega] \mapsto I_{\mathscr{Y}}[\omega], \tag{2}$$

du groupe $H^m\mathscr{X}$ dans $\mathbb{R}$ (i.e. une fonctionnelle linéaire sur $H^m\mathscr{X}$)

Le nombre $I_{\mathscr{Y}}[\omega]$ s'appelle *période* de la forme ω (ou de la classe de cohomologie $[\omega]$) sur la sous-variété $\mathscr{Y}$.

Il est évident que si pour la classe de cohomologie $[\omega]$ il existe une sous-variété $\mathscr{Y}$ telle que $I_{\mathscr{Y}}[\omega] \neq 0$, alors $[\omega] \neq 0$. Il se trouve que la réciproque — qui est un théorème très difficile ! — est vraie aussi : *si* $[\omega] \neq 0$, *il existe une sous-variété* $\mathscr{Y} \subset \mathscr{X}$ *telle que* $I_{\mathscr{Y}}[\omega] \neq 0$.

La difficulté de ce théorème s'explique par le fait qu'en construisant la sous-variété $\mathscr{Y}$ à l'aide de la forme ω il est très difficile de remplir toutes les conditions auxquelles doit satisfaire cette sous-variété (par exemple l'absence de self-intersections). Il est clair à priori que la démonstration devient plus simple si l'on affaiblit ces conditions, i.e. si l'on élargit la classe des sous-variétés $\mathscr{Y}$ (ou, plus exactement, des fonctionnelles (2)).

La première idée qui vient en tête est de remplacer les sous-variétés $\mathscr{Y}$ par des *sous-variétés singulières* fermées, i.e. par des applications différentiables de la forme $\gamma : \mathscr{D} \to \mathscr{X}$, où $\mathscr{D}$ est une variété

orientée compacte sans bord à m dimensions. L'autre idée consiste à envisager des combinaisons linéaires arbitraires de fonctionnelles (2). La collation de ces deux idées nous conduit à des fonctionnelles de la forme

$$[\omega] \mapsto \sum_{i=1}^{N} a_i \int_{\gamma_i} \omega, \quad [\omega] \in H^m \mathscr{X},$$

où γ_i sont des applications différentiables de la forme $\mathscr{D}_i \to \mathscr{X}$, et a_i des nombres réels quelconques (les variétés $\mathscr{D}_i$ étant fermées). Du reste il nous sera commode pour des raisons techniques évidentes d'introduire des combinaisons linéaires formelles de la forme

$$\gamma = \sum_{i=1}^{N} a_i \gamma_i \tag{3}$$

et de convenir que

$$\int_{\gamma} \omega = \sum_{i=1}^{N} a_i \int_{\gamma_i} \omega.$$

On pourrait s'attendre que pour la classe de cohomologie $[\omega] \neq 0$ il soit plus facile de construire une combinaison linéaire γ telle que $\int_{\gamma} \omega \neq 0$ que de trouver une sous-variété $\mathscr{Y}$ telle que $\int_{\mathscr{Y}} \omega \neq 0$. (Une fois que la combinaison γ est trouvée, on peut au besoin envisager de chercher $\mathscr{Y}$.)

Cette attente est en fait justifiée mais malheureusement pas dans la mesure voulue, car le problème reste encore assez compliqué. D'où la nécessité d'une nouvelle idée.

Cette idée fut proposée au début du siècle par l'illustre mathématicien français H. Poincaré et jeta les bases de la topologie algébrique moderne.

Poincaré proposa de considérer des combinaisons linéaires (3) de variétés singulières à m dimensions $\gamma_i : \mathscr{D}_i \to \mathscr{X}$ pour lesquelles les variétés $\mathscr{D}_i$, supposées comme précédemment compactes (pour que toutes les intégrales existent), puissent avoir un bord. La frontière d'une telle combinaison linéaire γ est définie par la formule

$$\partial\gamma = \sum_{i=1}^{N} a_i \, \partial\gamma_i,$$

où, rappelons-le, $\partial\gamma_i = \gamma_i |_{\partial\mathscr{D}_i}$. (Il est évident que si toutes les variétés $\mathscr{D}_i$ sont fermées, alors $\partial\gamma = 0$, mais l'égalité $\partial\gamma = 0$ peut aussi avoir lieu lorsque les variétés $\mathscr{D}_i$ sont à bord.) Il est clair par linéarité que *la formule de Stokes reste valable dans ce cas aussi*, i.e.

pour toute forme ω de degré $m-1$ sur $\mathscr{X}$ on a

$$\int_{\gamma} d\omega = \int_{\partial\gamma} \omega. \tag{4}$$

Donc, si $\partial\gamma = 0$, la formule

$$I_{\gamma}[\omega] = \int_{\gamma} \omega, \quad [\omega] \in H^m\mathscr{X},$$

définit de façon unique une fonctionnelle linéaire

$$I_{\gamma} : H^m\mathscr{X} \to \mathbb{R},$$

et le théorème exprime que pour toute classe de cohomologie $[\omega] \neq 0$, il existe une combinaison linéaire (3) (avec $\partial\gamma = 0$) telle que $I_{\gamma}[\omega] \neq 0$.

Il s'avère que sous cette forme le théorème se prouve assez aisément, car il est facile de trouver les variétés singulières correspondantes $\gamma_i : \mathscr{D}_i \to \mathscr{X}$ dans une classe de variétés élémentaires telles que les variétés $\mathscr{D}_i$ soient difféomorphes à une boule. [Ceci s'explique bien au niveau de l'intuition : chaque variété $\mathscr{D}_i$ peut être découpée en parties élémentaires difféomorphes à une boule. et l'intégrale d'une forme quelconque étendue à $\mathscr{D}_i$ est égale à la somme des intégrales étendues à ces parties.] Or s'il en va ainsi, on peut faire le pas suivant proposé aussi par Poincaré et n'envisager dès le départ que des variétés singulières $\mathscr{D} \to \mathscr{X}$ pour lesquelles la variété $\mathscr{D}$ est une boule. Les constructions correspondantes n'impliquent que des intégrations sur des boules, donc on peut les effectuer à un niveau élémentaire sans faire intervenir la théorie générale d'intégration sur des variétés.

Ceci étant, il est commode d'introduire, à la place des boules des cubes ou des simplexes (ni le cube ni le simplexe ne sont de toute évidence des variétés à bord, mais les points anguleux forment un ensemble négligeable et sont sans effet sur les intégrales). Nous choisirons — surtout par tradition — des simplexes (bien que de toute évidence la réduction des intégrales multiples à des intégrales itérées soit plus simple à réaliser pour les cubes).

* * *

Définition 1. On appelle *simplexe standard* Δ^m *à m dimensions* un sous-ensemble de l'espace $\mathbb{R}^{m+1}$ composé des points $\mathbf{t} = (t_0, t_1, \ldots, t_m)$ pour lesquels

$$0 \leqslant t_0 \leqslant 1, \quad 0 \leqslant t_1 \leqslant 1, \ldots, 0 \leqslant t_m \leqslant 1$$

et

$$t_0 + t_1 + \ldots + t_m = 1.$$

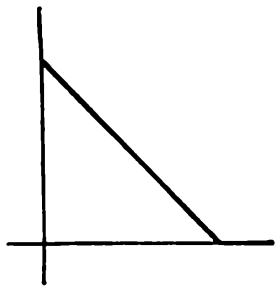
[C'est un simplexe euclidien au sens de la leçon 21 à sommets aux extrémités des vecteurs unitaires $\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_m$ de la base canonique de l'espace $\mathbb{R}^{m+1}$.]

Pour $0 \leqslant i \leqslant m$ $(m > 0)$ est définie l'application

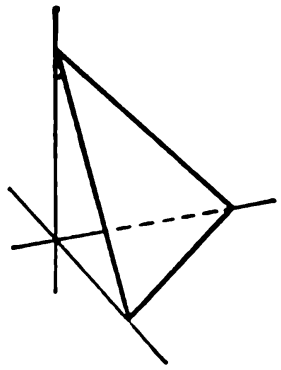
$$\delta_i : \Delta^{m-1} \to \Delta^m,$$

$$(t_0, t_1, \ldots, t_{m-1}) \mapsto (t_0, t_1, \ldots, t_{i-1}, 0, t_i, \ldots, t_{m-1}),$$

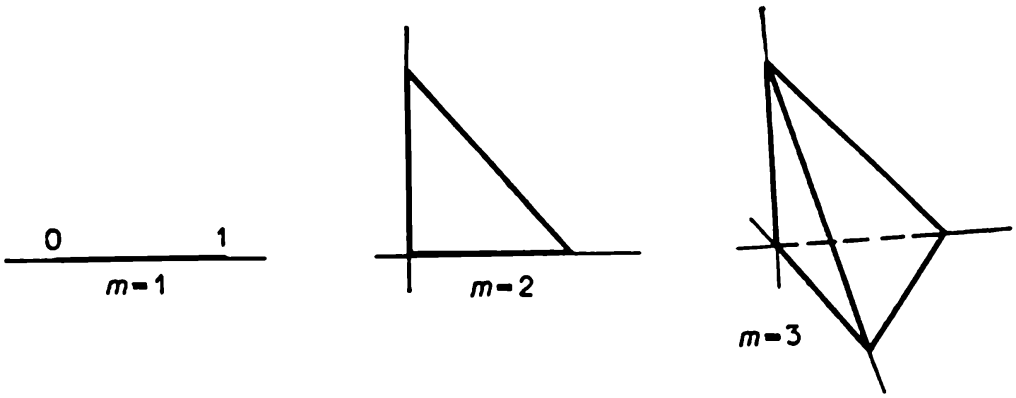
dont l'image (appelée généralement *i-ième face* du simplexe Δ^m) est composée de tous les points $\mathbf{t} \in \Delta^m$ tels que $t_i = 0$.



Simplexe standard Δ^1



Simplexe standard Δ^2



Simplexes standards projetés sur $\mathbb{R}^m$

Chaque point du simplexe Δ^m est défini de façon unique par ses coordonnées $t_1, \ldots, t_m$, ce qui permet d'identifier Δ^m au sous-ensemble de l'espace $\mathbb{R}^m$ composé des points $\mathbf{t} = (t_1, \ldots, t_m)$ tels que

$$0 \leqslant t_1 \leqslant 1, \ldots, 0 \leqslant t_m \leqslant 1$$

et

$$0 \leqslant t_1 + \ldots + t_m \leqslant 1$$

(i.e. avec le simplexe de l'espace $\mathbb{R}^m$ de sommets $\mathbf{0}, \mathbf{e}_1, \ldots, \mathbf{e}_m$).

On dit qu'une application $K \to \mathscr{X}$ d'un ensemble $K \subset \mathbb{R}^m$ dans une variété différentiable $\mathscr{X}$ est *différentiable* si elle est la restriction à K d'une application différentiable $U \to \mathscr{X}$, où U est un sous-ensemble ouvert de l'espace $\mathbb{R}^m$ contenant K.

On peut en particulier parler des applications différentiables de Δ^m dans $\mathscr{X}$.

Définition 2. On appelle *simplexe singulier à m dimensions* d'une variété différentiable $\mathscr{X}$ une application différentiable $\sigma : \Delta^m \to \mathscr{X}$. L'application (différentiable aussi de toute évidence)

$$\sigma \circ \delta_i : \Delta^{m-1} \to \mathscr{X}$$

est *sa i-ième face* $\partial_i \sigma$, $0 \leqslant i \leqslant m$. L'ensemble de tous les simplexes singuliers à m dimensions de la variété $\mathscr{X}$ sera désigné par le symbole $S_m \mathscr{X}$.

Définition 3. On appelle *chaîne singulière à m dimensions d'une variété $\mathscr{X}$ sur un groupe G* (ou *à coefficients dans G*) une combinaison linéaire formelle finie

$$\gamma = \sum_{\sigma \in S_m \mathscr{X}} a_\sigma \sigma, \quad a_\sigma \in G, \tag{5}$$

de complexes singuliers $\sigma \in S_m \mathscr{X}$ à coefficients dans le groupe G. (La finitude de la somme (5) exprime que seul un nombre fini de coefficients a_σ sont non nuls.) Ces chaînes forment un groupe par rapport à l'addition que nous désignerons par $C_m(\mathscr{X}; G)$. Pour $G = \mathbb{Z}$ ce n'est autre que le groupe abélien libre engendré par l'ensemble $S_m \mathscr{X}$, et pour $G = \mathbb{R}$, un espace vectoriel de base $S_m \mathscr{X}$.

On étudiera aussi les combinaisons linéaires (5) i n f i n i e s jouissant de la propriété suivante (appelée *propriété de finitude locale*; cf. définition 1 de la leçon 22) : pour tout point $p \in \mathscr{X}$ — donc pour tout ensemble compact $C \subset \mathscr{X}$ — il existe un voisinage U de p (de C) tel que l'ensemble des simplexes singuliers $\sigma \in S_m \mathscr{X}$, pour lesquels on a simultanément $a_\sigma \neq 0$ et $\sigma(\Delta^m) \cap U \neq \varnothing$, est fini. De telles combinaisons linéaires seront appelées *chaînes infinies*. Elles forment un groupe $C_m^{\inf}(\mathscr{X}; G)$ dont $C_m(\mathscr{X}; G)$ est un sous-groupe.

Il est clair que *l'égalité $C_m(\mathscr{X}; G) = C_m^{\inf}(\mathscr{X}; G)$ a lieu si et seulement si la variété $\mathscr{X}$ est compacte.*

Pour tout $m \geqslant 1$ définissons les applications homomorphes

$$\partial : C_m(\mathscr{X}; G) \to C_{m-1}(\mathscr{X}; G)$$

et

$$\partial : C_m^{\inf}(\mathscr{X}; G) \to C_{m-1}^{\inf}(\mathscr{X}; G)$$

en admettant que pour toute chaîne (5) finie (ou resp. infinie)

$$\partial \gamma = \sum_{\sigma \in S_m \mathscr{X}} a_\sigma \partial \sigma. \tag{6}$$

où

$$\partial\sigma = \sum_{i=0}^{m} (-1)^i \partial_i \sigma,$$

i.e.

$$\partial\gamma = \sum_{\tau \in S_{m-1}\mathscr{X}} \left(\sum_{i=0}^{m} \sum_{\substack{\sigma \in S_m\mathscr{X}, \\ \partial_i\sigma = \tau}} (-1)^i a_\sigma \right) \tau,$$

où la sommation intérieure est étendue à tous les simplexes $\sigma \in S_m\mathscr{X}$ tels que $\partial_i\sigma = \tau$ (l'ensemble des simplexes tels que $a_\sigma \neq 0$ est fini — même si la chaîne γ est infinie — et par suite, cette somme a un sens).

Définition 4. La chaîne $\partial\gamma$ s'appelle *frontière* de la chaîne γ. Les chaînes telles que $\partial\gamma = 0$ s'appellent *cycles*

* * *

Pour toute forme différentielle ω de degré m sur une variété $\mathscr{X}$ et tout simplexe singulier $\sigma : \Delta^m \to \mathscr{X}$ est définie sur Δ^m (plus exactement sur un ensemble ouvert $U \supset \Delta^m$) la forme $\sigma^*\omega$. En coordonnées $t_1, \ldots, t_m$ cette forme s'écrit

$$\sigma^*\omega = w\, dt_1 \wedge \ldots \wedge dt_m,$$

où $w = w(\mathrm{t})$ est une fonction différentiable. Posons par définition

$$\int_\sigma \omega = \int \ldots \int_{\Delta^m} w(\mathrm{t})\, d\mathrm{t}. \tag{7}$$

En d'autres termes,

$$\int_\sigma \omega = \int_0^1 dt_1 \int_0^{s_1} dt_2 \ldots \int_0^{s_m} w(\mathrm{t})\, dt_m, \tag{8}$$

où $s_1 = 1 - t_1, \ldots, s_m = 1 - t_1 - \ldots - t_m$.

Définissons par additivité l'intégrale de ω le long d'une chaîne γ à m dimensions à coefficients dans $\mathbb{R}$: si $\gamma = \sum a_\sigma \sigma$, alors

$$\int_\gamma \omega = \sum_{\sigma \in S_m\mathscr{X}} a_\sigma \int_\sigma \omega. \tag{9}$$

Ceci étant, si la chaîne γ est infinie, pour que la somme (9) ait un sens (soit finie), il faut exiger que la forme ω soit à support borné.

Théorème 1 (théorème de Stokes pour intégrales le long de chaînes). *Pour toute chaîne singulière γ à m dimensions d'une variété différentiable $\mathscr{X}$ à coefficients dans $\mathbb{R}$ et toute forme différentielle ω de*

degré $m-1$ (à support borné si γ est infinie), on a l'égalité

$$\int_{\gamma} d\omega = \int_{\partial\gamma} \omega. \tag{10}$$

[Cette égalité est certes un cas particulier de l'égalité générale (4), mais nous en donnons ici une démonstration directe basée sur les seules formules (6), (7) et (9).]

D é m o n s t r a t i o n. Des formules (6) et (9) il s'ensuit immédiatement qu'il suffit de prouver l'égalité (10) dans le cas où la chaîne γ est un simplexe σ. Ceci étant, sans perdre en généralité on peut admettre que la forme $\sigma^*\omega$ s'écrit

$$\sigma^*\omega = w\, dt_1 \wedge \ldots \wedge \widehat{dt_k} \wedge \ldots \wedge dt_m,$$

où $0 \leqslant k \leqslant m$ (et le signe $\widehat{\ }$ indique comme toujours que le facteur chapeauté doit être omis), donc que la forme $\sigma^*(d\omega) = d(\sigma^*\omega)$ s'écrit

$$\sigma^*(d\omega) = (-1)^{k-1} \frac{\partial w}{\partial t_k}\, dt_1 \wedge \ldots \wedge dt_m.$$

Alors

$$\int_{\sigma} d\omega = \int_{\Delta^m} \ldots \int (-1)^{k-1} \frac{\partial w}{\partial t_k}\, dt =$$

$$= (-1)^{k-1} \int_{\Delta^{m-1}_{(k)}} \ldots \int \left(\int_0^{1-t} \frac{\partial w}{\partial t_k}\, dt_k \right) dt_{(k)},$$

où $t = t_1 + \ldots + \hat{t}_k + \ldots + t_m$, $dt_{(k)} = dt_1 \ldots \widehat{dt_k} \ldots dt_m$, et $\Delta^{m-1}_{(k)}$ est le simplexe Δ^{m-1} dans l'espace $\mathbb{R}^{m-1}$ muni des coordonnées $t_1, \ldots, \hat{t}_k, \ldots, t_m$. Puisque

$$\int_0^{1-t} \frac{\partial w}{\partial t_k}\, dt_k = w \Big|_{t_k=0}^{t_k=1-t} = w|_{t_k=1-t} - w|_{t_k=0},$$

où $w|_{t_k=r}$ est la restriction de la fonction w à l'hyperplan $t_k = r$ (traitée comme une fonction de $t_1, \ldots, \hat{t}_k, \ldots, t_m$), ceci prouve que

$$\int_{\sigma} d\omega = (-1)^{k-1} \int_{\Delta^{m-1}_{(k)}} \ldots \int (w|_{t_k=1-t} - w|_{t_k=0})\, dt_{(k)}. \tag{11}$$

Par ailleurs, puisque $t_i \circ \delta_i = 0$ (et donc que $\delta_i^*(dt_i) = 0$), il s'ensuit que $(\partial_i\sigma)^* \omega = \delta_i^*(\sigma^*\omega) = 0$ pour $i \neq 0$, k et, par suite,

$$\int_{\partial\sigma} \omega = \int_{\partial_0\sigma} \omega + (-1)^k \int_{\partial_k\sigma} \omega.$$

Dans le même temps, comme

$$t_i \circ \delta_k = \begin{cases} t_i & \text{si } i < k, \\ 0 & \text{si } i = k, \\ t_{i-1} & \text{si } i > k, \end{cases}$$

où au premier membre $t_1, \ldots, t_m$ sont des coordonnées dans $\mathbb{R}^m$ et au second membre $t_1, \ldots, t_{m-1}$, des coordonnées dans $\mathbb{R}^{m-1}$, il vient

$$(\partial_k\sigma)^* \omega = \delta_k^*(\sigma^*\omega) = (w \circ \delta_k)\, dt_1 \wedge \ldots \wedge dt_{m-1},$$

où

$$(w \circ \delta_k)(t_1, \ldots, t_{m-1}) = w(t_1, \ldots, t_{k-1}, 0, t_k, \ldots, t_{m-1}),$$

et donc

$$\int_{\partial_k\sigma} \omega = \int \ldots \int_{\Delta^{m-1}} w(t_1, \ldots, t_{k-1}, 0, t_k, \ldots, t_{m-1})\, dt_1 \ldots dt_{m-1}.$$

En désignant les variables $t_k, \ldots, t_{m-1}$ par $t_{k+1}, \ldots, t_m$, on peut mettre cette intégrale sous la forme suivante:

$$\int_{\partial_k\sigma} \omega = \int \ldots \int_{\Delta^m_{(k)}} (w|_{t_k=0})\, dt_{(k)}.$$

De façon analogue, puisque $t_i \circ \sigma_0 = t_{i-1}$, $i = 1, \ldots, m$, où $t_0 = 1 - t_1 - \ldots - t_{m-1}$, on a

$$\int_{\partial_0\sigma} \omega = \int \ldots \int_{\Delta^{m-1}} w(1 - t_1 - \ldots - t_{m-1}, t_1, \ldots, t_{m-1})\, dt_1 \ldots dt_{m-1}.$$

En faisant le changement de variables

$$\begin{array}{ll}
t'_1 = 1 - t_1 - \ldots - t_{m-1}, & t_1 = t'_2, \\
t'_2 = t_1, & t_2 = t'_3, \\
\ldots\ldots\ldots\ldots\ldots\ldots & \\
t'_{k-1} = t_{k-2}, & t_{k-2} = t'_{k-1}, \\
t'_{k+1} = t_k, & t_{k-1} = 1 - t'_1 - \ldots - t'_k - \ldots - t_m, \\
t'_{k+2} = t_{k+1}, & t_k = t'_{k+1}, \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\
t'_m = t_{m-1} & t_{m-1} = t'_m
\end{array}$$

(dont le jacobien est égal à $(-1)^{k-1}$) et en éliminant les accents, on trouve que

$$\int_{\partial_0\sigma} \omega = (-1)^{k-1} \int \ldots \int_{\Delta_{(k)}^{m-1}} (w|_{t_k=1-t})\, dt_{(k)}.$$

Donc,

$$\int_{\partial\sigma} \omega = (-1)^{k-1} \int \ldots \int_{\Delta_{(k)}^{m-1}} (w|_{t_k=1-t} - w|_{t_k=0})\, dt_{(k)}. \tag{12}$$

La collation des formules (11) et (12) prouve le théorème. □

Du théorème 1 il résulte que pour tout cycle γ l'intégrale

$$I_\gamma \omega = \int_\gamma \omega$$

de la forme fermée ω le long de γ dépend seulement de la classe de cohomologie $[\omega] \in H^m\mathscr{X}$ de cette forme, i.e. la formule

$$I_\gamma [\omega] = \int_\gamma \omega \tag{13}$$

définit de façon unique un homomorphisme (une fonctionnelle linéaire)

$$I_\gamma : H^m\mathscr{X} \to \mathbb{R}. \tag{14}$$

De façon analogue pour tout cycle γ infini, la même formule (14) envisagée avec une forme ω à support borné définit de façon unique une fonctionnelle linéaire

$$I_\gamma : H^m_{\text{s.b.}}\mathscr{X} \to \mathbb{R}. \tag{15}$$

Dans les deux cas le nombre $I_\gamma [\omega]$ s'appelle *période* de la forme ω (ou de la classe de cohomologie $[\omega]$) suivant le cycle γ.

* * *

Etudions maintenant la dépendance de la fonctionnelle I_γ par rapport au cycle γ. Pour cela il nous faut développer préalablement l'appareil algébrique correspondant.

Lemme 1. *Si* $0 \leqslant i < j \leqslant m$, *alors*

$$\partial_i \partial_j \sigma = \partial_{j-1} \partial_i \sigma \tag{16}$$

pour tout simplexe singulier σ *à* m *dimensions de la variété* $\mathscr{X}$.

D é m o n s t r a t i o n. Puisque $\partial_i \partial_j \sigma = \sigma \circ \delta_j \circ \delta_i$ et $\partial_{j-1}\partial_i \sigma = \sigma \circ \delta_i \circ \delta_{j-1}$, il suffit de prouver que

$$\delta_j \circ \delta_i = \delta_i \circ \delta_{j-1}, \quad 0 \leqslant i < j \leqslant m.$$

Mais par définition l'action de l'application δ_i consiste à mettre un zéro à la i-ième place dans le vecteur $\mathbf{t} = (t_0, \ldots, t_m)$. Puisque après cette opération le numéro $j > i$ aura la place qu'occupait le numéro $j - 1$, l'application $\delta_i \circ \delta_{j-1}$ mettra un zéro à la i-ième et à la j-ième place. D'autre part, après l'action de l'application δ_j, toutes les places de numéros inférieurs à j garderont le leur et donc l'application $\delta_j \circ \delta_i$ mettra aussi un zéro à la i-ième et à la j-ième place. Par conséquent, $\delta_j \circ \delta_i = \delta_i \circ \delta_{j-1}$. □

Proposition 1. *Pour toute chaîne γ à m dimensions ($m \geqslant 2$) d'une variété $\mathscr{X}$ sur un groupe G on a l'égalité*

$$\partial\partial\gamma = 0. \tag{17}$$

Démonstration. (Comparer avec la démonstration de la proposition 1 de la leçon 21.) Il suffit de toute évidence de prouver l'égalité (17) pour le cas où la chaîne γ est un simplexe singulier $\sigma \in S_m\mathscr{X}$. Or, pour tout simplexe σ la chaîne $\partial\partial\sigma$ est visiblement la somme de chaînes de la forme $(-1)^{i+j}\partial_i\partial_j\sigma$, où $0 \leqslant i \leqslant m-1$ et $0 \leqslant j \leqslant m$. D'autre part, en vertu du lemme 1, si $i < j$, on a $\partial_i\partial_j\sigma = \partial_{j-1}\partial_i\sigma$. Donc, pour tout couple (i, j) tel que $i < j$, on aura deux termes égaux dans la chaîne $\partial\partial\sigma$, l'un de signe $(-1)^{i+j}$, l'autre de signe opposé $(-1)^{(j-1)+i}$. Donc, les termes de la chaîne $\partial\partial\sigma$ se détruisent tous deux à deux. □

La proposition 1 exprime que pour tout $m \geqslant 1$ le *groupe de frontières* $B_m(\mathscr{X}; G)$ (l'image de l'homomorphisme $\partial : C_{m+1}(\mathscr{X}; G) \to C_m(\mathscr{X}; G)$) est contenu dans le *groupe de cycles* $Z_m(\mathscr{X}; G)$ (le noyau de l'homomorphisme $\partial : C_m(\mathscr{X}; G) \to C_{m-1}(\mathscr{X}; G)$), et de façon analogue, le *groupe de frontières infinies* $B_m^{\inf}(\mathscr{X}; G)$ (l'image de l'homomorphisme $\partial : C_{m+1}^{\inf}(\mathscr{X}; G) \to C_m^{\inf}(\mathscr{X}; G)$) est contenu dans le *groupe de cycles infinis* $Z_m^{\inf}(\mathscr{X}; G)$ (le noyau de l'homomorphisme $\partial : C_m^{\inf}(\mathscr{X}; G) \to C_{m-1}^{\inf}(\mathscr{X}; G)$). Donc, sont définis les groupes quotients

$$H_m(\mathscr{X}; G) = Z_m(\mathscr{X}; G)/B_m(\mathscr{X}; G)$$

et

$$H_m^{\inf}(\mathscr{X}; G) = Z_m^{\inf}(\mathscr{X}; G)/B_m^{\inf}(\mathscr{X}; G),$$

appelés *m-ième groupe d'homologie singulière* (resp. *d'homologie infinie singulière*) *de la variété $\mathscr{X}$ à coefficients dans le groupe G* (ou *sur le groupe G*).

Remarque 1. Les groupes d'homologie s'appellent aussi *groupes d'homologie à supports compacts*, et les groupes d'homologie infinie, *groupes d'homologie à supports arbitraires.*

Remarque 2. Cette construction s'étend intégralement à tout espace topologique $\mathscr{X}$ (à ceci près qu'il faut remplacer la différentiabilité par la continuité dans la définition d'un simplexe singulier). Pour une variété différentiable, outre les groupes d'homologie défi-

nis ci-dessus (qui sont dits basés sur des simplexes singuliers différentiables) il apparaît *ipso facto* des *groupes d'homologie topologiques* (basés sur des simplexes singuliers continus quelconques). On démontre cependant par une construction un peu laborieuse mais d'un principe élémentaire que les *seconds groupes s'appliquent naturellement sur les premiers par un isomorphisme.*

Les éléments des groupes d'homologie (les classes des groupes de cycles suivant les groupes de frontières) s'appellent *classes d'homologie* et deux cycles d'une même classe d'homologie (i.e. différant par leur frontière) seront dits *homologues.* Respectivement, les frontières seront appelées *cycles homologues à* 0.

* * *

Supposons maintenant que le groupe G est de nouveau le corps $\mathbb{R}$.

On sait que pour un cycle γ et une forme fermée ω l'intégrale

$$\int_\gamma \omega$$

dépend de la seule classe de cohomologie $\mathbf{x} = [\omega]$ de la forme ω. Du théorème 1 il résulte maintenant que cette intégrale dépend aussi de la seule classe ξ d'homologie du cycle γ (car

$$\int_{\partial\beta} \omega = \int_\beta d\omega = 0$$

pour toute chaîne β). Donc, la formule

$$\langle \xi, \mathbf{x} \rangle = \int_\gamma \omega \tag{18}$$

définit de façon unique un *couplage* (cf. leçon II.4) des espaces vectoriels $H_m(\mathscr{X}; \mathbb{R})$ et $H^m\mathscr{X}$ (ainsi que des espaces vectoriels $H_m^{\inf}(\mathscr{X}; \mathbb{R})$ et $H^m_{\text{s.b.}}\mathscr{X}$).

On rappelle (*ibidem*) qu'un couplage d'espaces vectoriels $\mathscr{T}$ et $\mathscr{W}$ est dit *non dégénéré* si pour tout vecteur $\xi \in \mathscr{T}$ non nul, il existe un vecteur (automatiquement non nul) $\mathbf{x} \in \mathscr{W}$ tel que $\langle \xi, \mathbf{x} \rangle \neq 0$ et réciproquement.

Théorème 2. *Pour toute variété $\mathscr{X}$ orientée, paracompacte et séparée, les deux couplages* (18) *(de $H_m(\mathscr{X}; \mathbb{R})$ et $H^m\mathscr{X}$ d'une part, et de $H_m^{\inf}(\mathscr{X}; \mathbb{R})$ et $H^m_{\text{s.b.}}\mathscr{X}$ d'autre part) sont non dégénérés.*

Discutons ce théorème en détail. (Nous glisserons sur la démonstration.)

Pour chaque couplage des espaces vectoriels $\mathscr{T}$ et $\mathscr{W}$ (sur un corps $\mathbb{K}$) et tout vecteur $\xi \in \mathscr{T}$ la formule

$$I_\xi \mathbf{x} = \langle \xi, \mathbf{x} \rangle, \quad \mathbf{x} \in \mathscr{W},$$

définit une fonctionnelle linéaire $I_\xi : \mathscr{W} \to \mathbb{K}$, i. e. un vecteur de l'espace dual $\mathscr{W}' = \mathrm{Hom}\,(\mathscr{W}, \mathbb{K})$. La non-dégénérescence en $\mathbf{x}$ du couplage exprime que pour tout élément $\mathbf{x} \in \mathscr{W}$ non nul il existe un élément $\xi \in \mathscr{V}$ tel que $I_\xi \mathbf{x} \neq 0$.

De façon analogue, pour tout élément $\mathbf{x} \in \mathscr{W}$ la formule $I_\mathbf{x}\xi = \langle \xi, \mathbf{x} \rangle$ définit une fonctionnelle $I_\mathbf{x} : \mathscr{V} \to \mathbb{K}$, et la non-dégénérescence en ξ du couplage exprime que pour tout élément $\xi \in \mathscr{V}$ non nul, il existe un élément $\mathbf{x} \in \mathscr{W}$ tel que $I_\mathbf{x}\xi \neq 0$.

Pour le couplage (18) la fonctionnelle I_ξ est exactement la fonctionnelle I_γ (où γ est un cycle tel que $\xi = [\gamma]$). Donc, le théorème 2 affirme que pour toute forme ω fermée non cohomologue à 0 il existe un cycle γ (visiblement non homologue à 0) tel que $I_\gamma\omega \neq 0$. C'est exactement l'assertion par la discussion de laquelle nous avons entamé cette leçon.

Mais le théorème 2 affirme aussi que, réciproquement, pour tout cycle γ non homologue à 0, il existe une forme fermée ω (visiblement non cohomologue à 0) telle que $I_\omega\gamma \neq 0$ (et donc $I_\gamma\omega \neq 0$).

Le théorème 2 affirme par ailleurs que les formes à support borné et les chaînes infinies sont justiciables des mêmes assertions. [A v e r t i s s e m e n t. Contrairement à ce qu'on pourrait penser, ces deux parties du théorème ne se ramènent pas directement l'une à l'autre car une forme à support borné, finiment non cohomologue à zéro, peut être néanmoins cohomologue à zéro, et de façon analogue, une chaîne finie non homologue à zéro peut être la frontière d'une chaîne infinie.]

Les correspondances $\xi \mapsto I_\xi$ et $\mathbf{x} \mapsto I_\mathbf{x}$ définissent de toute évidence des homomorphismes

$$\mathscr{V} \to \mathscr{W}' \quad \text{et} \quad \mathscr{W} \to \mathscr{V}' \tag{19}$$

des espaces vectoriels $\mathscr{V}$ et $\mathscr{W}$ dans les espaces duaux $\mathscr{W}'$ et $\mathscr{V}'$. Ceci étant, dire que le couplage est non dégénéré revient à dire que *les homomorphismes* (19) *sont des monomorphismes.*

S'agissant des espaces finidimensionnels $\mathscr{V}$ et $\mathscr{W}$, on sait (proposition 4 de la leçon II.4) que pour un couplage non dégénéré, les applications (19) sont même des isomorphismes. Pour les espaces infinidimensionnels, ce n'est généralement pas le cas et il est même possible que l'une de ces applications soit un isomorphisme et l'autre non.

Exemple 1. Soient S un ensemble infini, $\mathscr{V}$ l'espace vectoriel des sommes finies formelles $\xi = \sum a_\sigma \sigma$, où $\sigma \in S$, $a_\sigma \in \mathbb{K}$, et $\mathscr{W}$ l'espace vectoriel des fonctions $\mathbf{x} : S \to \mathbb{K}$. La formule

$$\langle \xi, \mathbf{x} \rangle = \sum_{\sigma \in S} a_\sigma \mathbf{x}(\sigma)$$

définit alors de façon unique un couplage (visiblement, non dégénéré) de $\mathscr{V}$ et $\mathscr{W}$. Ce couplage est tel que chaque fonctionnelle

$$I_\xi(\mathrm{x}) = \sum_{\sigma \in S} a_\sigma \mathrm{x}(\sigma)$$

est telle que $I_\xi \mathrm{x} \neq 0$ pour les seuls éléments $\mathrm{x} \in \mathscr{W}$ d'un sous-espace finidimensionnel de l'espace $\mathscr{W}$ (composé des fonctions x non nulles sur les seuls éléments $\sigma \in S$ pour lesquels $a_\sigma \neq 0$). Donc, l'application $\xi \mapsto I_\xi$ de $\mathscr{V}$ dans $\mathscr{W}'$ n'est visiblement pas un épimorphisme. Au contraire, l'application $\mathrm{x} \mapsto I_\mathrm{x}$ de $\mathscr{W}$ dans $\mathscr{V}'$ est un isomorphisme, puisque toute fonctionnelle $\varphi : \mathscr{V} \to \mathbb{K}$ peut être représentée sous la forme I_x, $\mathrm{x} \in \mathscr{W}$ (il suffit de poser $\mathrm{x}(\sigma) = \varphi(\sigma)$ pour tout $\sigma \in S$).

Par conséquent, le théorème suivant précise le théorème 2 :

Théorème 3. *Pour toute variété séparée paracompacte $\mathscr{X}$ les homomorphismes*

$$H^m\mathscr{X} \to H_m(\mathscr{X}; \mathbb{R})', \quad \mathrm{x} \mapsto I_\mathrm{x}, \tag{20}$$

et

$$H_m^{\inf}(\mathscr{X}; \mathbb{R}) \to (H^m_{\text{s.b.}}\mathscr{X})', \quad \xi \mapsto I_\xi, \tag{21}$$

où I_x et I_ξ sont des fonctionnelles $H_m(\mathscr{X}; \mathbb{R}) \to \mathbb{R}$ et $H^m_{\text{s.b.}}\mathscr{X} \to \mathbb{R}$, définies par les formules

$$I_\mathrm{x}\xi = \int_\gamma \omega, \quad \mathrm{x} = [\omega], \quad \xi = [\gamma],$$

et

$$I_\xi \mathrm{x} = \int_\gamma \omega, \quad \mathrm{x} = [\omega]_{\text{s.b.}}, \quad \xi = [\gamma]^{\inf},$$

sont des isomorphismes.

[Au contraire, les applications « duales »

$$H_m(\mathscr{X}; \mathbb{R}) \to (H^m\mathscr{X})' \quad \text{et} \quad H^m_{\text{s.b.}}\mathscr{X} \to H_m^{\inf}(\mathscr{X}; \mathbb{R})'$$

ne sont généralement pas des isomorphismes.]

Le théorème 3 (combiné au théorème 2) est connu sous le nom de théorème de de Rham (qui le premier l'a prouvé pour les variétés compactes).

L'isomorphisme (20) ramène le calcul du groupe $H^m\mathscr{X}$ à celui des groupes $H_m(\mathscr{X}; \mathbb{R})'$, calcul pour lequel des méthodes spéciales, en principe suffisamment efficaces, ont été élaborées en topologie algébrique.

* * *

La construction des groupes d'homologie $H_m(\mathscr{X}; G)$ et $H_m^{\inf}(\mathscr{X}; G)$ se fait en deux étapes: la construction des groupes de chaînes $C_m(\mathscr{X}; G)$ et $C_m^{\inf}(\mathscr{X}; G)$ avec les homomorphismes ∂ (première éta-

pe) et la construction des groupes d'homologie à l'aide des premiers (deuxième étape).

Définition 5. La famille

$$C_\cdot : C_0 \xleftarrow{\partial} C_1 \leftarrow \ldots \leftarrow C_{m-1} \xleftarrow{\partial} C_m \leftarrow \ldots$$

de groupes et d'homomorphismes s'appelle *complexe de chaînes* si

$$\partial\partial\gamma = 0$$

pour tout élément $\gamma \in C_m$, $m \geqslant 1$. Les éléments du groupe C_m s'appellent *chaînes à m dimensions*; l'homomorphisme ∂, *opérateur bord*; les chaînes γ telles que $\partial\gamma = 0$, i.e. appartenant au noyau

$$Z_m C_\cdot = \operatorname{Ker}(\partial : C_m \to C_{m-1})$$

de l'homomorphisme $\partial : C_m \to C_{m-1}$, *cycles* (pour $m = 0$ on convient que $Z_m C_\cdot = C_0$); les chaînes de la forme $\partial\gamma$, i.e. appartenant à l'image

$$B_m C_\cdot = \operatorname{Im}(\partial : C_{m+1} \to C_m)$$

de l'homomorphisme $\partial : C_{m+1} \to C_m$, *frontières* (ou *cycles homologues à zéro*); le groupe quotient

$$H_m C_\cdot = Z_m(C_\cdot)/B_m(C_\cdot)$$

m-ième groupe d'homologie du complexe $C_\cdot$; ses éléments *classes d'homologie à m dimensions*, et les cycles d'une même classe d'homologie, i.e. différant par leur frontière, *cycles homologues.*

Les groupes $H_m(\mathscr{X}; G)$ sont donc les groupes d'homologie $H_m C_\cdot(\mathscr{X}; G)$ du complexe

$$C_\cdot(\mathscr{X}; G) = \{C_m(\mathscr{X}; G); \partial\},$$

et les groupes $H_m^{\inf}(\mathscr{X}; G)$, groupes d'homologie du complexe

$$C_\cdot^{\inf}(\mathscr{X}; G) = \{C_m^{\inf}(\mathscr{X}; G); \partial\}.$$

Lorsque le complexe de chaînes $C_\cdot = \{C_m; \partial\}$ est composé d'espaces vectoriels sur le corps $\mathbb{K}$, on peut construire pour tout $m \geqslant 0$ l'espace vectoriel dual

$$C^m = (C_m)' = \operatorname{Hom}(C_m; \mathbb{K}).$$

En mathématique moderne le passage au cas dual est traditionnellement signalé par le préfixe « co » (exemple: vecteur et covecteur). Par conséquent, les éléments de l'espace C^m s'appellent *cochaînes à m dimensions* du complexe de chaînes $C_\cdot$.

Pour toute cochaîne $c \in C^m$, la formule

$$(\delta c)(\gamma) = c(\partial\gamma), \quad \gamma \in C_{m+1},$$

définit une cochaîne δc et puisque

$$(\delta\delta c)(\gamma) = c(\partial\partial\gamma) = 0$$

pour toute chaîne $\gamma \in C_{m+2}$, on a $\delta\delta c = 0$ pour toute cochaîne $c \in C^m$. Par définition (cf. définition 1 de la leçon 20) cela signifie que *la famille $C^\cdot = \{C^m; \partial\}$ de groupes et d'homomorphismes est un complexe de cochaînes.*

Les cocycles $c \in C^m(C^\cdot)$ de ce complexe sont caractérisés par la relation

$$(\delta c)(\gamma) = c(\partial\gamma) = 0, \quad \gamma \in C_m,$$

i.e. par le fait qu'ils sont nuls sur le sous-groupe des frontières $B_m C_\cdot \subset C_m$. Par définition (cf. leçon II.4) ceci signifie que

$$Z^m C^\cdot = \operatorname{Ann} B_m C_\cdot$$

(*les cocycles forment l'annulateur du groupe des frontières*). De façon analogue, si $\partial\gamma = 0$, alors $(\delta e)(\gamma) = e(\partial\gamma) = 0$ pour toute cochaîne e, et inversement, si $c(\gamma) = 0$ lorsque $\partial\gamma = 0$, la formule $e(\partial\gamma) = c(\gamma)$ définit de façon unique une fonctionnelle linéaire $e : B_{m-1}C_\cdot \to \mathbb{K}$ telle qu'en la prolongeant arbitrairement — en préservant sa linéarité — à l'espace C_{m-1} tout entier, on obtient une cochaîne $e : C_{m-1} \to \mathbb{K}$ telle que $(\delta e)(\gamma) = e(\partial\gamma) = c(\gamma)$, i.e. telle que $\delta e = c$. Ceci prouve que

$$B^m C^\cdot = \operatorname{Ann} Z_m C$$

(*les cobords forment l'annulateur du groupe des cycles*).

La prolongeabilité de la fonctionnelle e du sous-espace $B_{m-1}C^\cdot$ à l'espace C_{m-1} tout entier est un fait général : pour tout sous-espace $\mathcal{P}$ d'un espace vectoriel $\mathcal{V}$ quelconque — voire même infinidimensionnel — chaque fonctionnelle linéaire $e : \mathcal{P} \to \mathbb{K}$ se prolonge à $\mathcal{V}$. Pour le prouver, il suffit de considérer l'espace complémentaire $\mathcal{Q}$ de $\mathcal{P}$, dont l'existence a été établie à la leçon 22 à la suite de la remarque 3, et de se donner une fonctionnelle arbitraire e sur $\mathcal{Q}$.

Pour les groupes de cohomologie $H^m(C^\cdot)$ du complexe $C^\cdot$ il en résulte que pour toute classe de cohomologie $\mathbf{x} \in H^m C^\cdot$, la formule

$$(\# \mathbf{x})(\xi) = c(\gamma), \quad \xi \in H_m C_\cdot,$$

où c est un cocycle arbitraire de la classe de cohomologie $\mathbf{x}$, et γ un cycle arbitraire de la classe d'homologie ξ, définit de façon unique une fonctionnelle

$$\# \mathbf{x} : H_m C_\cdot \to \mathbb{K},$$

et l'application

$$\# : H^m C^\cdot \to (H_m C_\cdot)', \quad \mathbf{x} \mapsto \# \mathbf{x},$$

obtenue est un isomorphisme. [La monomorphie de cette application résulte de l'égalité $B^m C^\cdot = \operatorname{Ann} Z_m C_\cdot$, quant à l'épimorphie, elle

est assurée par la possibilité de prolonger toute fonctionnelle linéaire $Z_m C_. \to \mathbb{K}$ à l'espace C_m tout entier.] Donc, *pour tout complexe de chaînes $C_.$ composé d'espaces vectoriels et tout $m \geqslant 0$ on a l'isomorphisme*

$$\# : H^m C^\cdot \to (H_m C_.)'.$$

Le calcul des groupes de cohomologie se ramène ainsi à celui des groupes d'homologie.

* * *

Pour le complexe $C_. = C_. (\mathscr{X}; \mathbb{K})$ des chaînes singulières de la variété différentiable $\mathscr{X}$ sur le corps $\mathbb{K}$ le complexe de cochaînes $C^\cdot$ est désigné par le symbole $C^\cdot (\mathscr{X}; \mathbb{K})$. Puisque pour tout $m \geqslant 0$ l'ensemble $S_m \mathscr{X}$ est une base de l'espace $C_m (\mathscr{X}; \mathbb{K})$ et que les fonctionnelles linéaires sur un espace vectoriel arbitraire s'identifient de façon naturelle aux fonctions à valeurs dans $\mathbb{K}$ définies sur sa base, la composante $C^m (\mathscr{X}; \mathbb{K}) = C_m (\mathscr{X}; \mathbb{K})'$ du complexe $C^\cdot (\mathscr{X}; \mathbb{K})$ s'identifie pour tout $m \geqslant 0$ à l'espace vectoriel des fonctions $S_m \mathscr{X} \to \mathbb{K}$. Ceci permet en particulier de définir le groupe $C^m (\mathscr{X}; G)$ pour un groupe G a r b i t r a i r e. Par définition, les éléments de ce groupe, qui sont appelés *cochaînes singulières à m dimensions de la variété $\mathscr{X}$ sur le groupe G*, sont des fonctions c à valeurs dans G définies sur l'ensemble $S_m \mathscr{X}$ des simplexes singuliers à m dimensions de la variété $\mathscr{X}$. L'opérateur cobord δ est défini par la formule

$$(\delta c)(\sigma) = \sum_{i=0}^{m} (-1)^i c(\partial_i \sigma)$$

qui détaille la formule $(\delta c)(\sigma) = c(\partial c)$.

Cocycles, cobords et classes de cohomologie du complexe $C^\cdot (\mathscr{X}; G)$ s'appellent *cocycles, cobords et classes de cohomologie singuliers* de la variété $\mathscr{X}$, et les groupes respectifs sont désignés par $Z^m (\mathscr{X}; G)$, $B^m (\mathscr{X}; G)$ et $H^m (\mathscr{X}; G)$. D'après le résultat général prouvé plus haut, pour tout corps $\mathbb{K}$ on a l'isomorphisme

$$H^m (\mathscr{X}; \mathbb{K}) \to H_m (\mathscr{X}; \mathbb{K})'. \tag{22}$$

Lorsque $\mathbb{K} = \mathbb{R}$, chaque forme $\omega \in \Omega^m \mathscr{X}$ définit à l'aide de la formule

$$\bar{\omega}(\gamma) = \int_\gamma \omega, \quad \gamma \in C_m (\mathscr{X}; \mathbb{R}),$$

une cochaîne $\bar{\omega} \in C^m (\mathscr{X}; \mathbb{R})$, et de plus

$$\overline{d\omega} = \delta \bar{\omega}. \tag{23}$$

[En effet

$$\overline{d\omega}(\gamma) = \int_{\gamma} d\omega = \int_{\partial\gamma} \omega = \overline{\omega}(\partial\gamma) = (\delta\overline{\omega})(\gamma)$$

pour toute chaîne $\gamma \in C_m(\mathscr{X}; \mathbb{R})$.] De la formule (23) il résulte que l'application $\omega \mapsto \overline{\omega}$ envoie cocycles dans cocycles, et cobords dans cobords. Donc, cette application induit un homomorphisme

$$H^m\mathscr{X} \to H^m(\mathscr{X}; \mathbb{R}). \tag{24}$$

La comparaison des définitions montre aussitôt que le composé de l'homomorphisme (24) et de l'isomorphisme (22) n'est autre que l'homomorphisme (20) du théorème de de Rham. Donc, pour prouver ce théorème (au moins pour les groupes $H^m\mathscr{X}$) il suffit d'établir que *l'homomorphisme* (24) *est un isomorphisme.* Nous n'avons malheureusement pas la possibilité de produire cette démonstration ici.

Remarque 3. L'assertion que l'homomorphisme (24) est un isomorphisme exprime que les deux approches de la définition des groupes de cohomologie d'une variété — en termes de formes et de chaînes singulières — conduisent en fait au même résultat. Il existe d'autres approches, par exemple celle de Čech-Leray de la leçon 22, qui, on le sait, conduit aux groupes isomorphes aux groupes $H^m\mathscr{X}$. Donc, pour prouver que (24) est un isomorphisme, il suffit de montrer que *les groupes de cohomologie singulière sont isomorphes aux groupes de Čech-Leray* (qui à la leçon 22 ont été désignés par le même symbole $H^m(\mathscr{X}; \mathbb{R})$). Cette approche qui ne fait pas intervenir les formes est valable pour tout espace topologique (pour lequel existent des recouvrements de Leray convenablement définis). On peut donc poser le problème de la légitimité de cette assertion sous sa forme générale (pour tout espace topologique). Il s'avère que la réponse est positive pour les polyèdres (cf. remarque 2 de la leçon 21) mais négative dans le cas général. L'étude de ces problèmes relève d'une théorie ramifiée qui fait partie de la topologie algébrique. Malheureusement nous n'avons pu en parler que par allusions. [Par exemple, la comparaison des définitions des groupes de cohomologie des schémas simpliciaux et des groupes de cohomologie singulière des variétés — et de façon plus générale des espaces topologiques — fait ressortir leur grande similitude. La mise en évidence de cette similitude dans les termes généraux conduit aux ensembles simpliciaux dont la théorie est élégante et profonde. D'autre part, la construction d'un complexe de cochaînes à l'aide d'un complexe de chaînes est un minable exemple et une pâle ombre des manipulations algébriques susceptibles d'être effectuées sur les complexes de chaînes en algèbre homologique, etc. Le lecteur intéressé pourra s'adresser pour d'autres renseignements à l'ouvrage de Botta et Thu cité dans l'Introduction ainsi qu'aux nombreux autres ouvrages de topologie algébrique.]

INDEX

Adhérence 127
Algèbre de Lie 166, 245
— — de matrices 246
Angle des courbes 55
Antidérivation 267
Application(s) de carte 95
— de classe C^r 42, 94
— de cochaînes 281
— continue 116
— de coordonnées 95
— différentiable 50, 94, 118, 385, 388, 414
— étale 192
— homotopes 372
— isométriques 58
— localement plane 195
— $F\mathcal{X}$-multilinéaire 264
— de passage 97
— proprement homotopes 372
— de rang maximal 198
— rétractante 132
Arcs simples 23
— — réguliers 26
Arête de rebroussement 48
Atlas 97
— de ∂-cartes 385
— équivalents 99
Axiome de dénombrabilité, deuxième 112
— —, premier 110

Base dénombrable 112
— d'ensembles ouverts 111
— mobile 67
— — de Frenet 32, 33, 37
— de voisinages 110
Bord 385
— de la sous-variété 391
Boule 19

Carte(s) 95, 385
— adaptée 378
— de bord 386
— centrées 180
— découpée 378
— intérieure 386
— positivement compatibles 355
— — orientées 356
— sécantes 95
Catégorie 117
Caténoïde 62
Chaîne(s) 153, 423
— infinies 414
— singulière 414
Chaînette 62
Champ(s) central 396
— de covecteurs 237
— irrotationnel 396
— potentiels 396
— solénoïdaux 396
— de tenseurs 235
— de vecteurs 236
Changement de cartes 96
— de coordonnées 44
— de paramètre 25
Chemin 172
— constant 172
— différentiable 176
— — par morceaux 176
— réciproque 172
Circulation 406
Classe(s) de cohomologie 281, 293, 307
— — à deux dimensions 298
— — singulières 425
— de différentiabilité 100, 386
— d'homologie 432
Cloison 134
Cobords 298, 306
— de degré m 281
— singuliers 425

Cochaîne(s) 305, 423
— singulières 425
Cocycles 306
— cohomologues 281, 293
— de degré m 281
— à deux dimensions 298
— à une dimension 292
— singuliers 425
Coefficients de connexité 85
— de la forme dans la carte 388
Coercition 143
Commutateur 244
Complexe acyclique 313
— de chaînes 423
— de cochaînes 281
— double 309
— — bordé 312
— de de Rham 281
Composante(s) connexe(s) 174
— — par arcs 173
— normale 401
— du vecteur dans la carte 180
— de l'unité 173
Coordonnées 43
— curvilignes 43
— gaussiennes 90
— locales 43, 93, 95
Coupe d'un sous-ensemble 218
Couplage 420
Courbe(s) 21, 247
— différentiable de classe C^r 23
— équivalentes 25
— fermée 392
— générique 31, 33, 37
— intégrale 247
— — maximale 248
— non paramétrées 25
— paramétrées 25
— régulière 24, 25
— simple 23, 25
Courbure(s) d'une courbe 30
— moyenne 73
— principales 73
— relative 31
— totale 73
Critère de convergence de Cauchy 162
Crochet de Lie 244
Cycle(s) 415
— de la différentielle 338
— homologues 420, 423

Degré d'une application 369, 377
Densité 349
— de la circulation 408
— différentiable 349
— localement bornée 349
— nulle 350
Densité presque continue 349
— des sources 404
— à support borné 350
— volumique 350
Dérivée de Lie 253
— normale 403
Déterminant de Gram 56
Diamètre d'un sous-ensemble 124
Difféomorphisme 52, 94, 118, 385
— $\mathbb{C}$-analytique 171
— local 192
Différentielle(s) de l'application 52, 189
— exactes 281
— extérieure 274
— de la suite 336
Dimension complexe 172
— de la variété 100, 385
Divergence 395
Domaine à frontière régulière 378

Elément(s) comparables 153
— maximal 153
Ensemble(s) bien ordonné 144
— centré 152
— connexe 14
— fermé 15
— inductif 153
— isomorphes 144
— maigre 215
— négligeable 219, 348
— ordonné 144
— ouvert 14
— partiellement ordonné 144
— séparés 137
— totalement ordonné 144
Equations différentielles sur la variété 248
— fonctionnellement indépendantes 206
— paramétriques de la courbe 22, 44
Espace(s) de Baire 215
— complètement normal 136
— connexe par arcs 173
— cotangent 191
— dénombrable à l'infini 342
— homéomorphes 110
— localement compact 342
— — euclidien 111, 215
— normal 136
— projectif réel 105
— régulier 215
— séparé 112
— tangent 200
— topologique 107
— — compact 123
— totalement discontinu 175

Exhaustion compacte 342
Extrémité(s) du chemin 172
— de la courbe 22
— d'une ligne 14

Famille localement finie 316
Filtration 327
Filtre 152
Flot 250
— maximal 250
Fonction(s) caractéristique 15, 348
— différentiable de classe C^∞ 23
— égales à proximité d'un point 239
— d'Urysohn 211
Forme(s) différentielle(s) de degré m 388
— — linéaires 258
— — sur la variété 263
— essentielle 363
— exactes 281
— fermées 281
— finiment cohomologue à zéro 363
— inessentielle 363
— normée 363
— quadratique, deuxième 70
— —, première 53
— —, troisième 71
— à support borné 388, 390, 391
Formule(s) de dérivation de Weingarten 85
— de Frenet 32, 34, 39
— de Gauss-Ostrogradski 383
— —, généralisée 409
— de Green 384, 404
— de Newton-Leibniz 384
— de passage 96
— de Peterson-Codazzi 87
Frontière de la chaîne 415

Gradient 191, 395
Graphique 14
Groupe(s) abélien gradué 326
— de cohomologie 281, 307
— — de Čech 319
— — du complexe double 311
— — de de Rham 281
— — du recouvrement 311
— différentiable 228
— de Lie 228
— — de matrices 166
— — — complexes 171
— symplectique 168
— — orthogonal 168
— — unitaire 170

Hélice 34
Hélicoïde 63
Homéomorphisme 94, 110
Homomorphismes bords 313
Homotopie 372
— différentiable 372
— propre 372
Hyperplan équatorial 103

Identité de Jacobi 245
Immersion 198
Indicatrice de Dupin 71
Intérieur d'un ensemble 109
— d'une ∂-variété 386
Isométrie 56, 58
Isomorphisme 144

Lemme de Zorn 154
Ligne de coordonnées 44
— fermée 14, 15
— ouverte 15
— semi-ouverte 15
— simple 14
Longueur d'arc 27

Matrice(s) jacobienne 189
— orthogonale 165
— propre 165
— pseudo-orthogonales 169
— symplectiques 168
— unimodulaire 165
— unitaires unimodulaires 170
Méridiens 45
Monéomorphisme 23, 117
Multiplicité d'un recouvrement 121

Nerf 303
Nombre de Betti 285
— de Lebesgue 124
Normale à la courbe 32

Opérateur cobord 306
— différentiel linéaire 238
— de Laplace 400
Ordinal 144
— dénombrable 145
— distingué 147
— fini 144
Origine du chemin 172
Orientation 356
— induite 356, 379

BeatriceGloria_personal library

Parallèles 46
Paramètre différentiel de Beltrami, mixte 401
— — —, premier 61
— — de Darboux 401
— naturel 28
Paramétrisation(s) 43
— équivalentes 44
— de la surface 43
Partition subordonnée 316
— de l'unité 316
Pavage de rang N 124
Période de la forme 410
Plan normal 37, 68
— osculateur 37
— rectifiant 37
— tangent 49
Poids local dénombrable 110
Point(s) adhérent 154
— affinement indépendants 303
— du bord 386
— critique 221
— elliptique 72
— hyperbolique 72
— intérieur 15, 93, 109, 386
— méplat 31
— non singulier 16
— parabolique 72
— positif 356
— régulier 16
— simple 14
Polyèdres 305
Potentiel 396
— newtonien 397
— vecteur 396
Produit des chemins 173
— direct 129, 151, 228, 317
— intérieur 265
Profil 45
Projection stéréographique 103
Pseudo-rayon 79
Pseudo-sphère 79
Puits 404

Rayon de courbure 31
Réalisation géométrique 304
Recouvrement coercible 143
— inscrit 343
— de Leray 318
— localement fini 343
— numérotable 316, 344
— ouvert 111, 123
Repère mobile de Frenet 32, 33
Réseau de coordonnées 44
Retract 132
— de voisinage 374
Rotationnel 395

Schéma simplicial 303
Section normale 68
Signe de l'application 37)
Simplexe(s) 304
— abstraits 304
— euclidiens 304
— ordonnés 305
— singulier 414
— standard 412
Source 404
Sous-ensemble clairsemé 215
— connexe 173
— cubable 349
— distingué 303
— fermé 15, 126
— nulle part dense 215
— ouvert 15
— ouvert-fermé 173
— ∂-ouvert 385
Sous-espace tangent 49
Sous-variété(s) immergée 199
— ouverte 109
— plongée 199, 388
— singulières 391, 410
Structure différentiable 99
— — standard 100, 158
— induite 109
— topologique 107
Submersion 198
Suite dégénérée 341
— spectrale 336
Support de la carte 95
— de la paramétrisation 43
Surface(s) des binormales 48
— à bord 390
— développables 65
— élémentaire 43
— isométriques 58
— des normales principales 48
— réglée 46
— de révolution 45
— des tangentes 48
Symboles de Christoffel 85
Système fondamental de voisinages 110

Tenseur antisymétrique 262
Théorème des applications étales 193
— d'approximation polynomiale 375
— de Brouwer 387
— de changement des variables 348
— de Dieudonné 347
— d'invariance homotope du degré 373
— de de Rham 320, 422
— de de Rham-Leray 319
— de Sard 221
— de Stokes 380, 389, 391, 415

Théorème de Tykhonov 152
— de Weierstrass 375
— de Whitney 17
Topologie 107
— induite 111
— de l'ordre 147
Torsion 33
Tractrice 79
Transfert 268

Ultrafiltre 153

Valeur régulière 204
Variété(s) analytiques 100, 172
— à bord 387
— sans bord 387
Variété(s) de classe C^r 100
— difféomorphes 119
— différentiable 99, 385
— fermées 387
— grassmannienne 161
— paracompacte 316
— parallélisable 237
— plongeables 209
— orientable 355, 388
— orientée 356, 358, 388
— de Stiefel 161
— topologique 100
— des vecteurs tangents 230
Vecteur binormal 33
— normal 67
— de la normale principale 33
— tangent 24, 49, 180, 186, 247
Voisinage de coordonnées 110